

Dd	541	gpfpgvsgvgfvgagdegdvrrrsipeliregpasggglpsfpssrpygdtaanaaka	500
Qy	601	kysaaavpvglgclgalgvgrvifsgwtwgaapaaaakaaakaaaoqlcvraaglcslgsvg	660
Dd	601	kysaaavpvglgclgalgvgrvifsgwtwgaapaaaakaaakaaaoqlcvraaglcslgsvg	660
Qy	661	glgyfvogvgigcippaaaaakaaakayagabgvlvgagagofplagvaarfgcglsrppgga	720
Dd	661	glgyfvogvgigcippaaaaakaaakayagabgvlvgagagofplagvaarfgcglsrppgga	720
Qy	721	cckacgcbkxk 731	
Dd	721	cckacgcbkxk 731	
RESULT 2			
R56853			
ID	R56853	standard; Protein; 733 AA.	
XX	XX	R56853;	
Dt	XX	22-MAR-1995 (first entry)	
DE	XX	Synthetic human tropoelastin (SHEL).	
KW	XX	Tropoelastin; pharmaceutical; surgical dressing.	
OS	XX	Synthetic.	
PX	XX	M09414958-A.	
PR	XX	07-JUL-1994.	
PF	XX	16-DEC-1993; 93KW-AU00655.	
PR	XX	22-DEC-1992; 92AU-0006520.	
PR	XX	28-JUN-1993; 93AU-0009661.	
PA	XX	(UNSY) UNIV STDNEY.	
DR	XX	Martin SL, Weiss AS;	
DX	XX	WPJ: 1994-26363/32.	
DX	XX	N-PSEB: Q70941.	
PT	XX	Synthetic polynucleotide(s) - encode recombinant tropoelastins	
PT	XX	and variants	
XS	XX	Disclosure; Page 30; 77pp; English.	
CC	XX	Human synthetic tropoelastin is susceptible to hydrolytic breakdown	
CC	XX	of the crosslinks. Such material may be useful in e.g. surgical	
CC	XX	applications, where the gradual loss of material over time is	
CC	XX	intended.	
SQ	XX	Sequence 733 AA:	
Query Match		100.0%; Score 3785; DB 15; Length 733;	
Local Similarity		100.0%; From 1a-213	
Matches 731;		Mismatches 0; Indels 0; Gaps	
Qy	1	GCVGFAIGDVGFGVFPYFCAGLGGALGGALGPSPKFKVPFGSLAGAGLGALGFAPVT	60
Dd	3	gvgvgaalggvgvgvyfpgaglgallggallgpggkpkvpvgslagaglgaglgafpvt	62
Qy	61	F6GALPVGVDAAAYAKAKAGAGLGGVFVGGLGV/SAGAVPFGACVFEGKVPGVL	120
Dd	63	fpgaivpgvvadaaaykaakagaglgvgvfyvgvglyvgaagvvpagvgvkgvpgvyl	122
Qy	121	PCVPGVGLVLCARFGVGVLFPGVTPGAVKPA-VGGAFAFG-FGVPPGPGPGCPGVFLY	180

141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

[illegible]

13-AUG-1987 (Rel. 05, Last sequence update)
 30-MAY-2000 (Rel. 39, Last annotation update)
 DEFINED BY: ELASTIN PROTEIN (PROPELASTIN).
 GN: ELASTIN
 OS: Bos taurus (Bovine).
 OC: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 OX: NCBI_TaxID=9913;
 [1]
 SEQUENCE FROM N.A.
 RP: MEDLINE=69194172; PubMed=3032943;
 RT: "Primary structures of bovine elastin a, b, and c deduced from the
 FT: sequences of cDNA clones.";
 J. Biol. Chem. 262:5755-5762(1987).
 [2]
 SEQUENCE OF 1-27 FROM N.A.
 RP: TIS552-Elastin Ligament;
 MEDLINE=69274159; PubMed=2543440;
 RT: "El H.; Anderson M.; Grinstein-Golstein N.; Bashir M.M.;
 FT: Rosenbloom J.C.; Adams W.K.; Iolik Z.; Yoon K.; Mecham R.,
 BA: "Structure of the bovine elastin gene and S1 nuclease analysis of
 FT: alternative splicing of elastin mRNA in the bovine nuchal ligament.";
 Biochemistry 28:2365-2370(1989).
 [3]
 SEQUENCE OF 1-27 FROM N.A.
 RP: MEDLINE=91234332; PubMed=2031719;
 RT: "Mandrich A.; Shi W.; Anwar R.A.;
 FT: Mandrich A.; Shi W.; Anwar R.A.;
 BA: "Primary structure of bovine elastin gene: comparison with the
 FT: gene for human elastin.";
 Biochem. Cell. Biol. 69:185-192(1991).
 [4]
 -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
 CC: NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
 CC: -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
 CC: INTO AN EXTENSIBLE 3D NETWORK.
 CC: -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
 CC: -1- SUPPLEMENTARY NOTES: ISOTOPES: A (SHOWN HERE), B AND C: ARE
 CC: ALTERNATIVE ELASTIN CHAINS.
 CC: -1- PMW: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
 [5]
 THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage By and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 [6]
 EMBL; J02717; AAA30503.1;
 DR: EMBL; K03503; AAA30505.1;
 DR: EMBL; K03506; AAA30506.1;
 DR: EMBL; J02855; AAA30776.1;
 DR: EMBL; M58652; AAA03519.1;
 DR: PIR; A36728; A36728.
 DR: PIR; E36728; E36728.
 DR: PIR; C56728; C56728.
 DR: SWISS-PROT; P02717;
 KW: Elastic Protein; Connective tissue; Repeat; Signal;
 Alternative splicing.
 FT: SIGNAL 1 26
 FT: CHAIN 27 747 ELASTIN.
 FT: MOD_RES 105 105 OXIDATIVE DEAMINATION.
 FT: MOD_RES 109 109 OXIDATIVE DEAMINATION.
 FT: MOD_RES 252 252 OXIDATIVE DEAMINATION.
 FT: MOD_RES 271 271 OXIDATIVE DEAMINATION.
 FT: MOD_RES 275 275 OXIDATIVE DEAMINATION.
 FT: MOD_RES 327 327 OXIDATIVE DEAMINATION.
 FT: MOD_RES 327 327 OXIDATIVE DEAMINATION.
 FT: MOD_RES 400 400 OXIDATIVE DEAMINATION.
 FT: MOD_RES 404 404 OXIDATIVE DEAMINATION.
 FT: MOD_RES 407 407 OXIDATIVE DEAMINATION.

RESULT	2				
ELS_BOVIN					
ID	ELS_BOVIN	STANDARD;	PRT;	747	AA.
AC	P04985;	P04986;	P04987;	Q29421;	
DT	13-AUG-1987	(rel. 05,	Created)		

GenCore version 4.5
Copyright (c) 1995 - 2000 CompuGen Ltd.
OM protein - protein search, using sw model
Run on: April 24, 2001, 16:35:14 ; Search time 44.88 seconds
(without alignments)
557.949 Million cell updates/sec
Title: US-09-340-736-1
Perfect score: 3785
Sequence: 1 GGVFGARFGVGGVYFGR.....LSPIFFGACLGACGKRRK 731
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
-ched: 93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters: 93435
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : SwissProt_39.*
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3575.5	94.5	730	1 ELS_HUMAN	P15502 homo sapien
2	2479	65.5	747	1 ELS_BOVIN	P04985 bos taurus
3	2387	63.1	860	1 ELS_MOUSE	P54320 mus musculus
4	2311	61.1	864	1 ELS_RAT	Q99372 rattus norv
5	1542.5	40.8	750	1 ELS_CHICK	P07916 gallus gall
6	719	19.0	5263	1 PPOW_BOMBO	P05790 bombyx mori
7	675.5	17.8	1786	1 CA24_CAREL	P17140 caenorhabdi
8	662	17.2	1791	1 CA24_CAREL	P17140 caenorhabdi
9	662	17.2	1791	1 CA24_CAREL	P17140 caenorhabdi
10	662	17.2	1791	1 CA24_CAREL	P17140 caenorhabdi
11	609.5	15.1	1659	1 CA14_HUMAN	P02462 homo sapien
12	599.5	15.6	1659	1 CA14_HUMAN	P02462 homo sapien
13	589.5	15.6	1659	1 CA14_MOUSE	P02463 mus musculus
14	580	15.3	778	1 XQ34_MYCTU	P71933 mycobacteri
15	580	15.3	1454	1 CA11_HUMAN	P02452 homo sapien
16	578	15.3	1450	1 CA11_CANFA	Q9XSJ7 canis fami
17	577.5	15.3	1453	1 CA11_CHICK	P02457 gallus gall
18	575.5	15.2	1453	1 CA11_HUMAN	P02457 gallus gall
19	572.5	15.1	1384	1 CA21_BOVIN	P02463 bos taurus
20	570.5	15.1	1356	1 CA21_BOVIN	Q45372 canis fami
21	570.5	15.1	1356	1 CA21_BOVIN	Q45372 canis fami
22	570.5	15.1	1356	1 CA21_BOVIN	Q45372 canis fami
23	570.5	15.1	1356	1 CA21_BOVIN	Q45372 canis fami
24	563	14.9	1418	1 CA12_HUMAN	P02458 homo sapien
25	560.5	14.8	1453	1 CA11_MOUSE	P11087 mus musculus
26	559	14.8	1373	1 CA21_MOUSE	P08121 mus musculus
27	556	14.7	1459	1 CA12_MOUSE	Q01149 mus musculus
28	553.5	14.6	1366	1 CA21_HUMAN	P28481 mus musculus
29	553	14.6	1712	1 CA24_HUMAN	P08123 homo sapien
30	551	14.6	1327	1 CA24_HUMAN	P08572 homo sapien
31	548.5	14.5	1327	1 CA24_HUMAN	P08572 homo sapien
32	548.5	14.5	1327	1 CA24_HUMAN	P08572 homo sapien
33	539.5	14.3	1758	1 CA14_CAREL	P17133 caenorhabdi

P46804 nephila cia
P05997 homo sapien
P08122 mus musculus
P08123 homo sapien
P02459 mycobacteri
P02459 mycobacteri
P30754 xifia pach
Q07092 homo sapien
P53420 homo sapien
P12107 homo sapien
P13942 homo sapien
Q61245 mus musculus
Q01955 homo sapien

ALIGNMENTS

RESULT 1	ELS_HUMAN	STANDARD:	PRG: 730 AA.
ID	EL5502:		
AC	01-APR-1990 (Rel. 14, Created)		
DT	01-APR-1990 (Rel. 14, Last sequence update)		
DE	ELASTIN PRECURSOR (TROPOLASTIN).		
GN	ELN. sapiens (human)		
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID:9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Skin fibroblast;		
EX	MEDLINE=89009560; PubMed=3171221;		
RA	Fazio M.J., Olsen D.R., Rauh E.A., Baldwin C.F., Indik Z.,		
RA	Ornstein-Goldstein N., Ieh H., Rosenbloom J., Utico J.; fibroblast		
RT	Cloning of human length elastin cDNA from a human skin		
RT	utilizing exon-specific oligonucleotides";		
RL	J. Invest. Dermatol. 91:458-464(1988).		
RN	[2]		
RP	SEQUENCE OF 503-730 FROM N.A.		
RC	TISSUE=Placenta, and Hippocampus;		
EX	MEDLINE=96291399; PubMed=8689688;		
RA	Frangiskakis J.M., Ewart A.K., Morris C.A., Mervis C.B.,		
RA	Bertrand J., Robinson B.F., Klein B.P., Busing G.J., Everett D.A.,		
RA	Green E.D., Proeschel C., Gutowski N.J., Noble M., Atkinson D.L.,		
RA	Goldberg S.J., Keating M.F.;		
RT	Altered length elastin chain structure and localization of alternative splicing		
RT	constructive coagulation";		
RL	Cell 86:59-69(1996).		
CC	-1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND		
CC	NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.		
CC	-1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER		
CC	INTO AN EXTENSIBLE 3D NETWORK.		
CC	-1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.		
CC	-1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.		
CC	-1- DISEASE: DELETED IN WILLIAMS-BEUREN SYNDROME (WBS). A		
CC	DEVELOPMENTAL DISORDER CAUSED BY HAPLOINSUFFICIENCY FOR GENES IN A		
CC	2-CM REGION OF CHROMOSOME BAND 7Q11.23. HEMIZYGOUS DELETION MAY		
CC	BE RESPONSIBLE TO CERTAIN DEFECTS SUCH AS HIPERCALCAEMIA AND GROWTH		
CC	DELAY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see http://www.isb-sib.ch/announcement/		
CC	or send an email to license@sib-ch).		
CC	-----		
CC	EMBL: X15603; CA333627.1;		
CC	EMBL: X36860; X36860.1;		
CC	EMBL: X36860; X36860.1;		

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:31:20 ; Search time 62.39 Seconds
(without alignments)
225.086 Million cell updates/sec

title: US-09-340-736-1
Perfect score: 3785
Sequence: 1 GVPFGAIPGVGVFGA.....LSPIFGACAGACGKRX 731

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Residues: 18210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA*
1: /cgm2_5/prodata/2/aaa/5A_COMB.pep.*
2: /cgm2_5/prodata/2/aaa/5B_COMB.pep.*
3: /cgm2_5/prodata/2/aaa/6A_COMB.pep.*
4: /cgm2_5/prodata/2/aaa/6B_COMB.pep.*
5: /cgm2_5/prodata/2/aaa/6C_COMB.pep.*
6: /cgm2_5/prodata/2/aaa/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3785	100.0	731	2	US-08-911-364-1
2	3747.5	100.0	732	2	US-08-678-033A-40
3	1159.5	30.6	986	1	US-08-212-237-5
4	1159.5	30.6	988	5	PCT-US96-02772-5
5	1159	30.6	832	1	US-08-212-237-4
6	1159	30.6	832	5	PCT-US95-02772-4
7	1130	29.9	936	1	US-08-212-237-3
8	1130	29.9	936	5	PCT-US95-02772-3
9	1126.5	29.8	877	1	US-08-397-633A-54
10	1123	29.7	884	1	US-08-435-641A-58
11	1123	29.7	884	2	US-08-435-641A-58
12	1123	29.7	884	2	US-08-707-237A-96
13	1123	29.7	884	5	PCT-US96-06229-15
14	1123	29.7	884	5	PCT-US96-06229-15
15	1122	29.6	1002	2	US-08-707-237A-103
16	1122	29.6	1002	3	US-08-642-246-25
17	1122	29.6	1002	5	PCT-US96-06229-25
18	1118	28.5	1413	1	US-08-175-155-39
19	1118	28.5	1413	2	US-08-707-237A-45
20	1118	28.5	1464	1	US-08-477-509B-74
21	1118	28.5	1464	3	US-08-482-085B-74
22	1116	28.5	1056	5	PCT-US95-02772-6
23	1116	28.5	1056	5	PCT-US95-02772-6
24	1107	28.2	2257	1	US-08-175-155-47
25	1107	28.2	2257	1	US-08-477-509B-82
26	1107	28.2	2257	2	US-08-707-237A-53
27	1107	28.2	2257	3	US-08-482-085B-82

Sequence 108, Appl
Sequence 30, Appl
Sequence 70, Appl
Sequence 7, Appl
Sequence 46, Appl
Sequence 81, Appl
Sequence 52, Appl
Sequence 81, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 34, Appl
Sequence 34, Appl
Sequence 72, Appl
Sequence 72, Appl
Sequence 48, Appl
Sequence 77, Appl
Sequence 28, Appl

ALIGNMENTS

US-08-911-364-1
Sequence 1: Application US/08011364
Patent No. 565105
GENERAL INFORMATION:
APPLICANT: ROTHSTEIN, Aser
APPLICANT: KEELY, Fred W.
TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN ELASTIN AND OTHER FIBROUS PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESS: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
COMPUTER: IBM PC compatible
CURRENT RELEASE: Patent Release #1.0, Version #1.30
CURRENT RELEASE DATE: 07-AUG-1997
APPLICATION NUMBER: US/08/911,364
FILING DATE: 07-AUG-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/023,552
FILING DATE: 07-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE NUMBER: 041082/0104
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 731 amino acids
TYPE: amino acid
STRAND: linear
MOLECULE TYPE: Peptide
US-08-911-364-1

Query Match 100.0%; Score 3785; DB 2; Length 731;
Best Local Similarity 100.0%; Pred. No. 56-246;
Matches 731; Conservative 0; Mismatches 0; Indels 0;

[illegible]

[illegible]

RESULT 11
 US-08-435-641-15
 : Sequence 15, Application 05/08435641
 : Patent No. 5817303
 : GENERAL INFORMATION:
 : APPLICANT: Stetronsky, Erwin R.
 : TITLE OF INVENTION: Tissue Adhesive Using Synthetic
 : : Casing Linings
 : NUMBER OF SEQUENCES: 15
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Flehr, Hohbach, Test, Albrighton & Herbert
 : STREET: Four Embarcadero Center, Suite 3400
 : CITY: San Francisco
 : STATE: CA

TITLE OF INVENTION: Crosslinking
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSES: Fleier, Hobbach, Test, Albritton & Herbert
 STREET: Four Embarcadero Center, Suite 3400
 CITY: San Francisco
 STATE: CA

[illegible]

```

DB      832 GK 833

RESULT 13
; US-08-642-246-15
; Sequence 15, Application US/08642246
; Patent No. 6033654
; GENERAL INFORMATION:
; APPLICANT: STEDRONSKY, Erwin R.
; INVENTOR: CAPPELLO, Joseph
; TITLE OF INVENTION: Tissue Adhesive Using Synthetic
; POLYMERIZATION AND CRYSTALLINITY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLHBR, HOEBACH, TEST, ALBRIGHTON & HERBERT
; STREET: Four Embarcadero Center, Suite 200
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM: disk
; MEDIA TYPE: floppy
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/642.246
; FILING DATE:
; CLASSIFICATION: 435
; ACTIVITY/CODING INFORMATION:
; NAME: HOWLAND, Bertalan I.
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: A61127-1,B/R
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 884 amino acids
; TYPE: amino acid
; STRATEGY USED: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-642-246-15

Query Match          29.7%; Score 1123; DB 3; Length 884;
Best Local Similarity 41.8%; Pred. No. 6.4e+68;
Matches 352; Conservative 53; Mismatches 249; Indels 188; Gaps

QY      QY              2 GVPG-AIFG-GVGF-----GVYFGAGLCALGGAGLGPS-----CGKPLPVPG-G 43
DB       DB             61 GVPKRGVGVPVGVPVGVPVGVPVGAGSAGASGAGSGAGSVGVPVGVPVG 120
QY      QY              44 LAGAIGLGA-----GLGAFPAVTFFPALVPG-VGADAAMAYAKAAAGLA----GVPGVG 93
DB       DB             121 VEGVPGVKGVPGVGY-PGVGYPGVGVGTGAGAGSAGASGAGSAGSAGSVGPVG 179
QY      QY              94 --GLVSAGAVPDCAGATFK--VEGVLGVPTVPGVLPGARFPVGVLFPGVTPGAKVK 150
DB       DB             180 VPGTGV-----RVGVY-RNKGVFTVTFPY---GV-FGVGTVTGAGASGSAGSGAG 226
QY      QY              151 PKARVGVGAPAGIRGVPGFGSPGVGLGVPPIKAPKLFGVGLVPTTKLFLPVGTGPGVA 210
DB       DB             227 AGSSAGAGSGSVGRVGVPGVPGVPGVY--GVF-----CGKVP-----GVGVPY- 267
QY      QY              211 GAAGKAGVPYGCVGVGAAAAAARAATAKTGAGAGLVPG--GAGVP--GVPG-AIFG 285
DB       DB             266 VGVG-VPGV--GAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSAGSVGPVG 321
QY      QY              266 IG-----GLACTVTPAAAARAAAKYKAGAGLVGEFGSGFPGVTPGAGVPGVP 378
DB       DB             326 VGEPVGVPGVPGVPGAGASGSAGSAGS--GAGAG--SSAGASGS--VGVPVGPVGP 378


```


US-08-707-237A-103

```

: Sequence 103, Application US/08707237A
: Patent No. 5830713
: GENERAL INFORMATION:
: APPLICANT: Ferrari, Franco A.
: APPLICANT: Capello, Joseph
: APPLICANT: Crisman, Joe W.
: APPLICANT: V. N.
: TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
: TITLE OF INVENTION: REPETITIVE DNA
: NUMBER OF SEQUENCES: 108
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Holbach, Test, Albrinton & Herbert
: STREET: Four Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER: IBM PC compatible
: MEDIA TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/707,237A
: FILING DATE: 03-SEP-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/175,155
: FILING DATE: 09-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/053,049
: FILING DATE: 22-APR-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/609,716
: FILING DATE: 06-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/269,429
: FILING DATE: 09-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/114,618
: FILING DATE: 26-OCT-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 06/927,258
: FILING DATE: 04-NOV-1986
: ATTORNEY/AGENT INFORMATION:
: NAME: Treccarlin, Richard F.
: REGISTRATION NUMBER: 31,801
: REFERENCE/DOCKET NUMBER: A-55186-10/WHB
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 398-3249
: TELEFAX: (415) 398-3249
: TELETEXT: 910 277799
: INFORMATION FOR SEQ ID NO: 103:
: SEQUENCE CHARACTERISTICS:
: TYPE: amino acid
: LENGTH: 1002 amino acids
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-707-237A-103

```

```

Query Match          29.68; Score 1122; DB 2; Length 1002;
Best Local Similarity 39.18; Pred. No. 8 se-68;
Matches 370; Conservative 54; Mismatches 271; Indels 252; Gaps 60;

```

```

QY 2 GVPG-AIPG-GVPG---GVYFGAGLGGGALGPG-----GKPLKVP 41
DB 63 GVPKGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVPVGVP 122
QY 42 G-----GLAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 87

```

```

DB 123 GVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 181
QY 88 GVPVG-----GLVGAGVVPVPGVPGVPGVPGVPGVPGVPGV 140
DB 182 GVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 240
QY 141 GVPFGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 195
DB 241 PGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 293
QY 196 TTGKLPGVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 233
DB 294 -----GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 346
QY 234 AKAAKAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 284
DB 347 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 405
QY 285 AKAAKAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 307
DB 407 A-----AGVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 461
QY 308 -VGVFGAGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 364
DB 462 GVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 521
QY 365 VG--GIETGVGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 407
DB 522 VGVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 580
QY 408 GVGISFGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 455
DB 581 GVGVPFGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 637
QY 456 ---APGVGLAPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 506
DB 638 GKGVPVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 694
QY 507 KV-----ANAGLAAAGLGGAGAGAGAGAGAGAGAGAGAGAG 546
DB 695 GVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 754
QY 547 V-GAGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 590
DB 755 VAGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 812
QY 591 PGALAAAGA-AKYGAGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 637
DB 813 PGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
QY 638 KAAKAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 688
DB 873 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 932
QY 689 LGGV-LGGAGGPFLLGVAAR-----PFGGLSPFFGAGAGAG 727
DB 933 VPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 978

```

```

Search completed: April 24, 2001, 16:36:20
Job time: 300 sec

```


It	Score	Query	Length	DB	ID	Description
1	3747.5	99.0	792	1	EABU	elastin precursor.
2	2603.5	68.8	770	2	S59623	tropoelastin - she
3	2488	65.7	747	1	EABO	elastin precursor.
4	2387	63.1	860	1	EAMS	elastin precursor
5	2311	61.1	864	1	EARF	elastin precursor
6	1696.5	44.8	784	2	A26601	elastin precursor
7	678.5	17.8	1758	2	A29350	hypothetical prote
8	675.5	17.8	1759	2	A29351	collagen alpha 2(I)
9	663.5	17.5	1591	2	A70806	hypothetical glyci
10	661.5	17.5	1592	2	A70807	hypothetical glyci
11	644.5	17.0	1763	2	S16368	major ampullate fi
12	635	16.8	2639	2	A31328	fibroin - Chloas
13	626	16.5	1489	2	A70807	hypothetical glyci
14	611	16.1	749	2	A70812	hypothetical glyci
15	609.5	16.1	1669	1	CGHU4B	collagen alpha 1(I)
16	599.5	15.8	1691	1	S32291	collagen alpha 5(I)
17	597.5	15.7	1306	2	A70534	hypothetical glyci
18	594.5	15.7	882	2	A70812	hypothetical glyci
19	591	15.6	1660	2	A70869	hypothetical glyci
20	589.5	15.6	1669	1	CGMS4B	collagen alpha 1(I)
21	583	15.4	1329	2	A70591	hypothetical glyci
22	582	15.4	1464	1	CGHUIS	collagen alpha 1(I)
23	580	15.3	1778	2	A70963	hypothetical glyci
24	577.5	15.2	1472	1	CGCHIS	collagen alpha 1(I)
25	574.5	15.2	1667	1	CGHUIS	hypothetical glyci
26	573	15.1	1767	2	A70895	collagen alpha 1(I)
27	570	15.1	1049	1	CGB075	collagen alpha 1(I)
28	567	15.0	783	2	A70524	collagen alpha 1(I)
29	567	15.0	1418	2	A74567	collagen alpha 1(I)

Arch. Biochem. Biophys. 241, 584-591, 1985

A>Title: Analysis of the 3' region of the sheep elastin gene.

A.Reference number: A24758; MOID:85305763

A.Molecule type: mRNA

A.Residues: 655-669,671-716,732-770 <XOD>

C:Superfamily: elastin

C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine

F:760-765/Disulfide bonds: #status predicted

Query Match 68.9%; Score 2603.5; DB 2; Length 770;
Best Local Similarity 67.9%; Pred. No.3.1e-113;
Matches 553; Conservative 26; Mismatches 81; Indels 155; Gaps 21;

QY 1 GVPGAIPGPGVPGVFPAGAGLAGGAGLPGGKLPKVPFGLAGLAGA-GLGAPAV 59
DB 27 GGVVLGVLGPGVPGVFPAGGLGGLGPGVSGVKGKVPGLVSGEAGSGCLPAG 86
QY 60 TPEGAL--VFQGVDAADAAK--AKAGN---GLGVGKPGVGLSGNAGVPGCA 148
DB 87 AFGPGFPGAGGAGAAAYAKAKAGAGGLGVGGVGGVGLSGVAGVPCPLGAGGV 106
QY 109 GVKPKVPVGLGVPVGGVLP--GARPVGVLPGVTFAGVGFAPGAGVAGFAGIPV 166
DB 147 GVKPKVPVGLGVPVGGVLPFGTGRFPGVLPVGTAGVKPAPGGGAFAGIPV 206
QY 167 GPFGPGVPGVPIGYPIAKPLPGGLTLPYTKLPYTGPGGVGAAGAGVPTGVSP 226
DB 207 GPVGGQPGVPLGIPIAKPLPGGLTLPYTKLPYTGPGGVGAAGAGVPTGVSP 266
QY 227 QAAALAAKAAKAAFGAGAGVLP--GVGAGVPGVGAIGTGLAGTGFAPAAANAAA 284
DB 267 Q-AAAAAKAAKLAGGAGVLPVGVGGAGIPGAFGAIPTGGIAGNAPDAANAAA 325
QY 285 AKAKTGAAGAGLVPGGPGVGVPGAGVPGVGVPGAGIP 327
DB 326 AKAKTGAAG--VFSPGVVPGAGVPGVPGVPGVPGVPGVPGVPGVPGV 378
QY 328 VPGAGIPGAAPVGVSPAPAAKAAKAAKAGVGVGGIGTPTGVGAGGPPG-- 382
DB 379 VPGVPIGVGVGAGVSPAAKAAKAAKAGVGVGGIPTEFGVPGGPGFGDAAQ 438
QY 383-----VVGVGITPVAG-VP-----SVGVPGVPGVPGVSGISPEAQAAAAKAYG 428
DB 439 APAKAAKAGVGGVGLGVLPGAPGAIPGVGGVP-----GV 479
QY 429 GPFPAAAKAAKAAQSLVPGVTPGVPGVPGVPGVPGVPGVPGVPGVPGV 486
DB 480 GAPAAAKAAKAAQGLGFGVGVARVGVPGVPGVPGVPGVPGVPGVPGV 538
QY 489 VAFGLSGGV--AAAAKAAKAAKAAQAAAGLGGATPGLGVPGVPGVAGVAGV 546
DB 539-----GLSPGGVTPAAKAAKAAKAAQFPAAAGLPAGVPGVPGVPGVAGVAGV 595
QY 547 VGAGVPGGAGGEGVRSILPELREGPSSSCLFSPSPSPFPGALAAKAAKAY 606
DB 596 AGA-----VFGLTAAKAAKAFKPGS 615
QY 607 PGVVLGSLGLGVGTPGVTVGSHAAAAKAAKAAKAAQGL--VGLAGLGGLG 663
DB 616 VGLAGVGLGSGATPVGSGVGFPAAKAAKAAKAAQGLGVLGVLGVLGVLG 675
QY 664 -VFGVGLGATPFAAAKAAKAAKAGLGGVLAGAGVPLGAGVAPPGVGLSP- 715
DB 676 AVFGAVLGVGSPFAAAKAAKAAKAGLGGVLAGAGVPLGAGVAPPGVGLSP 735
QY 716-----FPGNCIAGCAQGRKK 731
DB 736 LVGGCKPFPFGAGLALGFGPGCHLGSAGCKRRK 770

3

3

RESULT 3

Query Match 61.1%; Score 2387; DB 1; Length 860;
Best Local Similarity 61.1%; Pred. No. 3e-103;
Matches 546; Conservative 25; Mismatches 96; Indels 226; Gaps 34;
A:Map position: 5
C:Superfamily: elastin
F:1-72/words: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine
F:128-860/Product: elastin status predicted <Sig>
F:850-855/Disulfide bonds: status predicted <Mat>
F:854-855/Disulfide bonds: status predicted

Query Match 61.1%; Score 2387; DB 1; Length 860;
Best Local Similarity 61.1%; Pred. No. 3e-103;
Matches 546; Conservative 25; Mismatches 96; Indels 226; Gaps 34;
A:Map position: 5
C:Superfamily: elastin
F:1-72/words: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine
F:128-860/Product: elastin status predicted <Sig>
F:850-855/Disulfide bonds: status predicted <Mat>
F:854-855/Disulfide bonds: status predicted

Query Match 61.1%; Score 2387; DB 1; Length 860;
Best Local Similarity 61.1%; Pred. No. 3e-103;
Matches 546; Conservative 25; Mismatches 96; Indels 226; Gaps 34;
A:Map position: 5
C:Superfamily: elastin
F:1-72/words: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine
F:128-860/Product: elastin status predicted <Sig>
F:850-855/Disulfide bonds: status predicted <Mat>
F:854-855/Disulfide bonds: status predicted

Query Match 61.1%; Score 2387; DB 1; Length 860;
Best Local Similarity 61.1%; Pred. No. 3e-103;
Matches 546; Conservative 25; Mismatches 96; Indels 226; Gaps 34;
A:Map position: 5
C:Superfamily: elastin
F:1-72/words: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine
F:128-860/Product: elastin status predicted <Sig>
F:850-855/Disulfide bonds: status predicted <Mat>
F:854-855/Disulfide bonds: status predicted

```

|||||
Db 90 GAGLSY--ASRFGVGLVGGGAGAAAYAAKAAAGAGLGGIGVPGVGVGVGVGVGVGVGV 147
QY 93 -----GGGLYSAGAVPFGAGV-----KPGKVPVGLGVGVGVGVGVGVGVGVGV 134
Db 148 GYVGAGVGGIGGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 207
QY 135 PGVGLVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 193
Db 208 PGVGLVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 267
QY 194 PYTKGLKPYGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 253
Db 268 PYTKGLKPY-----GVGAGGKAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 321
QY 254 ASVPVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 313
Db 313 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 376
QY 376 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 357
Db 357 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 436
QY 358 GARGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 385
Db 437 GARGSV--GIPTVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 493
QY 386 -----GGIPGV--AGVP-----SVGVGVGVGVGVGVGVGVGVGVGVGVGV 427
Db 494 LVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 542
QY 428 VETPAARAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 487
Db 543 -----AAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 592
QY 488 GVAPG--TGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 546
Db 593 GLVPGDVGAGTFAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 650
QY 547 VGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 605
Db 651 AGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 679
QY 606 ---VGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 652
Db 680 GLGPGGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 738
QY 653 GL-----GGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 691
Db 739 GLGAGGLGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGG 798
QY 692 VLGAGGFPVLGVVAAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 724
Db 799 VL--GNRFFGVGVVAAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 857
QY 725 AGGRKK 731
Db 838 SCGRKK 864

RESULT 6
A:Residues: 1-212,237-524,535-784 <BR2>
A:Cross-references: GB:M15889; NTD:g212803; PIDN:AAA49108.1; PID:g212804
B:Baile, V.J.; Foster, J.A.
Biochem. Biophys. Res. Commun. 154, 1054-1060, 1988
A:Title: Multiple chick tropoelastin mRNAs.
A:Reference number: A30795; MUID:88309083
A:Accession: A30795
A:Molecule type: mRNA
A:Residues: 85-784 <BR>
A:Cross-references: GB:M21880; NTD:g212741; PIDN:AAA49082.1; PID:g212742
B:Okamoto, T.; Nishikawa, S.; Nishikawa, M.; Tajima, M.; Fukasawa, T.
Arch. Biochem. Biophys. 256, 455-461, 1987
A:Title: Sequence analysis of elastin cDNA from chick aorta and tissue-specific trans
A:Reference number: A27264; MUID:87297534
A:Accession: A27264
A:Molecule type: mRNA
A:Residues: 491-569, 'G', 571-604, 'A', 606-643, 'A', 645-687, 'R', 689-700, 'R', 702-784 <TOR>
A:Cross-references: GB:M18633; NTD:g211742; PIDN:AAA48761.1; PID:g211743
C:Superfamily: elastin
C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
F:1-24/domain: signal sequence (fragment) fstatus predicted <Sig>
F:237-524,535-784/Product: elastin fstatus predicted <AM>
F:773-779/Disulfide bonds: fstatus predicted

Query March 44.88; Score 1696.5; DB 2; Length 784;
Best Local Similarity 51.18; Pred. No. 1.5e-71;
Matches 450; Conservative 38; Mismatches 123; Indels 269; Gaps 54;
QY 1. GGVFGAIP--GGVPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 46
Db 25 GGVFGAIPGGVPGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGG 84
QY 47 --ALGAGLGGAPAVTFGLVPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 104
Db 85 GVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 138
QY 105 -----QPGAGV--KGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 150
Db 139 GGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 198
QY 151 PRAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 209
Db 199 AKGPG--AGAGLGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGGVGG 252
QY 210 AGA--AGKAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 268
Db 253 AGVLGAGKAGTGTGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 304
QY 269 I--AGVGTFAAAALAA--AAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 324
Db 305 VPGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 352
QY 325 IPVVG--AGTGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 382
Db 353 --LVPVGGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 404
QY 383 V--GVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 439
Db 405 VGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 453
QY 440 AKAAQGLV--VPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 497
Db 454 AKAAAGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 505
QY 498 VAAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 555
Db 506 V-----GVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 541
QY 555 AGAGGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 609
Db 542 PAA-----AAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 568
```

```

|||||
Db 90 GAGLSY--ASRFGVGLVGGGAGAAAYAAKAAAGAGLGGIGVPGVGVGVGVGVGVGV 147
QY 93 -----GGGLYSAGAVPFGAGV-----KPGKVPVGLGVGVGVGVGVGVGVGVGV 134
Db 148 GYVGAGVGGIGGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 207
QY 135 PGVGLVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 193
Db 208 PGVGLVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 267
QY 194 PYTKGLKPYGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 253
Db 268 PYTKGLKPY-----GVGAGGKAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 321
QY 254 ASVPVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 313
Db 313 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 376
QY 376 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 357
Db 357 GYVGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 436
QY 358 GARGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 385
Db 437 GARGSV--GIPTVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 493
QY 386 -----GGIPGV--AGVP-----SVGVGVGVGVGVGVGVGVGVGVGVGVGV 427
Db 494 LVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 542
QY 428 VETPAARAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 487
Db 543 -----AAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 592
QY 488 GVAPG--TGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 546
Db 593 GLVPGDVGAGTFAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 650
QY 547 VGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 605
Db 651 AGAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 679
QY 606 ---VGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 652
Db 680 GLGPGGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 738
QY 653 GL-----GGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 691
Db 739 GLGAGGLGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGGAGLGG 798
QY 692 VLGAGGFPVLGVVAAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 724
Db 799 VL--GNRFFGVGVVAAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 857
QY 725 AGGRKK 731
Db 838 SCGRKK 864

RESULT 6
A:Residues: 1-212,237-524,535-784 <BR2>
A:Cross-references: GB:M15889; NTD:g212803; PIDN:AAA49108.1; PID:g212804
B:Baile, V.J.; Foster, J.A.
Biochem. Biophys. Res. Commun. 154, 1054-1060, 1988
A:Title: Multiple chick tropoelastin mRNAs.
A:Reference number: A30795; MUID:88309083
A:Accession: A30795
A:Molecule type: mRNA
A:Residues: 85-784 <BR>
A:Cross-references: GB:M21880; NTD:g212741; PIDN:AAA49082.1; PID:g212742
B:Okamoto, T.; Nishikawa, S.; Nishikawa, M.; Tajima, M.; Fukasawa, T.
Arch. Biochem. Biophys. 256, 455-461, 1987
A:Title: Sequence analysis of elastin cDNA from chick aorta and tissue-specific trans
A:Reference number: A27264; MUID:87297534
A:Accession: A27264
A:Molecule type: mRNA
A:Residues: 491-569, 'G', 571-604, 'A', 606-643, 'A', 645-687, 'R', 689-700, 'R', 702-784 <TOR>
A:Cross-references: GB:M18633; NTD:g211742; PIDN:AAA48761.1; PID:g211743
C:Superfamily: elastin
C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
F:1-24/domain: signal sequence (fragment) fstatus predicted <Sig>
F:237-524,535-784/Product: elastin fstatus predicted <AM>
F:773-779/Disulfide bonds: fstatus predicted
```

QY	2	GVGP	-A	PGVPGVGVPG	---	ASGAGGSGGSGGSGKPLKPVPGGLAGAGG	-AGLGR	55
DB	634	GLTGGSGVP	PGPAG	PGYFEGKGGAGGLGSLG	---	XFGDGLGFLPGNKGSGAGVGGQPGVG	690	
QY	56	FWATFPGAL	---	VEGGYADAAATKAAKAGAGAGGAGGAGGAGGVLGVSAG	---	AV	102	
DB	631	FGKGGDGLGFLG	FTGQLQMGHGFAPEN	VNPGAPQGLGFLGFKTGKGGSGPGRGE	---	750		
QY	103	VPKAGGAGPKYKVEGWLGSV	---	YGVGSLPGAS	-FGVGVGLPVGYPTGAG	-VYKAP	154	
DB	751	VGCPRFPGFL	PGKMGDSGLGFPGLGPHG	-VPGDKGGVGGLGFLGPGKPGGVNDGFLP	---	808		
QY	155	GVGAGFAGT	PGVGPGRGG	---	QPGVP	---	LYGTAKPLPGSGTL	193
DB	809	GLNGO	-KSGFPGVPGGSGPFG	PLKDDAGLGLGFLGPGELSGRGGFP	-GAGLKGDSGL	866		
QY	194	PTTGTCKLPGYGGG	---	GVGAAAGKAGYPTGCV	---	224		
DB	867	PLGSGOFP	---	DGCKDAGLPGVPGPSSPGFP	PGDGLGCPVPGKGGDGLGFLGVTGL	923		
QY	225	---	GPQAAAAAAKAFKFGAGV	---	LFPGVGS	---	AGVPGCV	259
DB	924	KHNGAPQGS	GAGLPGAGGPGKGNAGT	PGVPGKGGDGLGFLGMLNGKPGGPGVM	---	983		
QY	280	PG	---	ATPGGAGV	---	GPAAAAAANKAKYKAGLVP	298	
DB	984	PCTPMKNGGGLG	FLGDSGLGVGPKMGKGGDFNGL	PGKGGAGCAAPDQKDGAG	-LIP	1041		
QY	259	GGSGF	-GP	-GVUGVGGAGVGS	---	VGPGAGLPGVPG	-AGTGGAGVGVVSPSEAAAKA	351

Query Match	17.88;	Score 675.5;	DB 2;	Length 1759;
Best Local Similarity	30.18;	Pred. No. 2.3e-24;		
Matches 284: Conservative	64;	Mismatches 323;	Indels 273;	Gaps 50;

QY	2	GVPK-AIRGSGVGSGVPGF-----AGTGLAGGGLGGGKFP-KVDPGGLAGAGGL-ASGLN	55
DB	635	GLRPGSPGVPGRGPGYFGKGGDGLFGLSG---KPSQDGLGFLGKNGKRGAGGPGQSPG	691
QY	56	FPATVTPGAL-----VPGVGVADAAAYAKAARGAGGVPVGVGGLGVSG-AY	102
DB	62	FGKAGDGLGFLGPGTGLQOMPFGVGNVNPAPGPGCPGLGFLGAGGVPGRPG	751
QY	103	VPQKGVGGKGVKPGVGLPGV-----YFGVGLGKAR-FPGVGVGLVFPVTPGAR-VKPAK	154

C:Species: Ascaris suum (pig roundworm)
CDate: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 13-Aug-1999
CAccession: SI6365
R:Petitt, J.; Kingston, I.B.
J. Biol. Chem. 266, 15149-15156, 1991
Articles: The complete primary structure of a nematode alpha-2(IV) collagen and the pa
CAccession: SI6366
CAccession: SI6366
CAccession: SI6366
Molecule type: mRNA
A:Residues: 1-1763 <BI>
A:Cross-references: GB:M67507; NID:q159648; PID:AAA18014.1; PID:q159649
C:Genetics:

C.x.y-zw: alternative splicing; basement membrane; cell binding; coiled coil; disulfide bond; extracellular matrix; integral component of membrane; protein complex; protein family; signal sequence; status predicted <SIG>
F.1-7/163/Product: collagen alpha 2(IV) chain status predicted <MAY>

F:27-193/ProDect: collagen alpha 2(IIV) chain #status predicted <NH>
F:27-42/Domain: non-collagenous NH1 #status predicted <NH>
F:43-1529/Domain: collagenous #status predicted <COL>
F:197-199/Region: cell attachment (R-G-D) motif.

```

F:1530-1763/Domain: carboxyl terminal noncatalytic; NCI status predicted <NC1>
F:1530-1638/Domain: repeat NC1 #status predicted <NC11>
F:1539-1763/Domain: repeat NC1 #status predicted <NC12>
F:31,34,39,41,536,539/Disulfide bonds: interchain #status predicted:

```

F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query: macch. 17.0%; Score 047; DB 2; Length 1703;
 Best Local Similarity 31.2%; Pred. No. 6.5e-23;
 Matches 277; Conservative 58; Mismatches 291; Indels 262; Gaps 54;

[illegible]

QY
2B GAAGGCGKPKFPG-----GLAGAGLGAGAGFAVTFPGALVPGGVADAATAKA 79

D6
485 GVRGPIGDSNLGPPGPRGGPQGNGVDGRDAPGLPGPKGDRGGTCAPCAHGAKG 545

Qy 80 AKAGAGLGGVPGVG---LGSAGAAVPP---QPGAGVKFGKTPGV---GLPGV--- 123

Dd 546 EKGDAGYAGLPGQPDGGLPGYCGATGAPGDDGLPGAPRGPG-PPPGQDGLPLPGGX 604

Qy
124 -----LPGGV-----LPGAR-----PFGVGTLPGVPFCAGVK----- 150

Dc
605 GEFDTLTINPGFPFGCKGKTGFPPFPGOGEGIFGKPGIVGAPGLPGFPFGKEPGLTGL 664

QY
151 PKAPVGGAFAATPGV-GFEGGQGVPGLP-IAK-K-LFGYGLPITTKSL----- 200

D6
665 PEKPCKDG-TPGLPLKGFGYCGQNP-GLPGMKDAGLPLGLPLPGAVGMGPVPKS 722

Qy 201 -----PCYCGGGVAGAAKACGPTGTGV-GPQAAAAAATAAKFGAGAGCVLPGV-G 252

Dd 723 QLRPPGGKDGLPLPGPKGEAGFPcAPGLQP-----AGLFG-LPGMKG 766

Qy 253 GAGCGVPG--ALPGLGGIACV-GTPAAAAAANAANKYGAAGL----- 296
 | : | | : | | : | : | : |
Db 767 NPLGPAQLAGLPGLPGKGIAGKPGLGLTGAKKGAG-YPGQPLPKPKGPSPWG 825

Qy 297 VPGGPG-----PGVVGPG-----AGVPG-----VGPG-AGIPWPGA-+*G 332
 |||| ||| ||| :||| ||| ||| +
 Db 826 PPGPPPPGLKGDIPGAPGIPLEGQGLPGVPGKGGLGLAGAGPGFCAKGEPP 885

Qy 333 TPGRVPGVVSPAAKAKAAYGARPVGVGIGITPTGYGAGFPG--FGVGVGGIPG 390
::: ||| :
Dd 886 LFG-LPGKEGPO-----GPQGQP---GAPGFPGCKGDEGLPGLPG 921

Q7 391 VAGVPSVGVPVGGVPGVISPEAQAARAAKAAKYGVGTAAAKAAKAAKAAQGLVPG 450
::: :|||: | | | : | | | | | |
Db 922 VSGMGDTLTPVPGVAG---PPGPGPFGO--KGDPFFSVAGAKGEA-----GL+GS-969

Q7 451 VGVAFGCVGVPGVGVPAGVGLAPGCVGVPAGVGPVAGVAPGICGPG-----GVAAAK 503

Db 970 LFGAPQKQEQGLAGLPGI--PCKMGAPGPGAGQDGLPGL-PGVKDRGKNGLEPK 1025
Qy 504 SNAKAAKAAQAAAGLGLAGTGLGAGVGVPGVGLGAGVGLGAGVGVPGFAGAGDEVR 563
Db 1026 GEPQAPAADEGEPEPL-PGQGLRGPQGPGL--PGLPGLKQDEQGPYGA----- 1074
Qy 564 RELSPELREDESSQHLPSRPSRVFVGAALAAKAAKAAKAAKAAKAAKAAKAAKAA 623
Db 1075 ----PGL-----WEGKGLGPGKXPGPAP-PQPLGDGPGFGLKNGPLGPGPLPG 1124
Qy 624 --GVVAGAPRAAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 670
Db 1125 QDGLPG----LPQKQESGFPQXGLVPGFGLPKNGKPGINGEKDAGLPLKRGGL 1179
Qy 671 GGIPFAAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 718
1180 DGL-PQKQAGPFGAPGPGVPGKGS---AGPFPPLKNGEPLPG 1223

RESULT 12
T31328 - chinese oak silkworm
C:Species: Anthrenus pernyi (Chinese oak silkworm)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T31328
R:Seizutsu, H.; Tamura, T.; Yukuhira, K.
submitted to the EMBL Data Library, August 1998
A:Description: Characterization of the full length fibroin gene of a wild silkworm, Anth.
A:Reference number: Z20995
A:Status: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2639 <SE>
A:Cross-references: EMBL:AF083334; NID:93450882; PID:93450883; PID:AC262606.1
A:JUNCTIONS: 14/3

Query Match 16.8%; Score 635; DB 2; Length 2639;
Best Local Similarity 29.8%; Pred. No. 2,3e-22;
Matches 258; Conservative 37; Mismatches 373; Indels 198; Gaps 25;
Qy 5 GAIPGPGVGVVY-----PQNGJLGGALGGALGGKPLKPLPVPSGL 44
1056 GSGAGVGGTGMGGDGYSDSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1115
45 AGA-----GLGGLGAGFAPVTPGALVPSG-----VADAAKAAKAAK 81
Db 1116 AAAPRAGHRAAGSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1175
Qy 82 AGAGLGGVGGVGLGVAGNVVPGAGVPGKVPKVPGLVPGVGLPGVGLGARFPGVGLP 141
Db 1176 AASGAGGGGGYGGYSDSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1228
Qy 142 GVTGAGVKNPAGVGNFAGIPGVPFGPGPVPVLPGLPIKAFKLPGLGGLPTTKMLP 201
Db 1229 AAAAAAAG 1283
Qy 202 YGVGPGVY-----AGAGAGAGVPTGVPVGPVGPVGPVGPVGPVGPVGPVGPVGP 234
Db 1284 YWGGDGYSDSAAAAAAG 1343
Qy 235 ----KAAKFGAGAGVLPVGGAG-----VFGVGAIPGTG 267
Db 1344 RAAGSAAAAAAG 1403
Qy 268 GIAGVGTFAAAAAAAG 317
Db 1404 GTGGTGGSDSAAAAAAG 1457
Qy 328 YFGAGPAGNATVPSPEAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 385

Db 1458 AAASGAGSGGGYGGYSDSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1508
Qy 386 GCTFVAGVPSGV 429
Db 1509 AAAAAAAG 1568
Qy 430 TFAAAAAAAG 489
Db 1569 SAAAG 1614
Qy 490 AFGVPGVGAAG 537
Db 1615 GTWGGGGYSDSAAAAAAG 1671
Qy 538 --VGAGVPGVGLGV 595
Db 1672 AAASGAGSGGGYGGYSDSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1731
Qy 595 KAKAA-----KYGAAPVGLGGLGAGVGPVGVGVGAGPA-----AAAKAAKAAK 642
Db 1732 AAAAARARAGHRAAGSAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1791
Qy 643 RAQPLGVAGAGLGVGLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 702
Db 1792 AAASGAGSGGGYGGY-----GVGGYSDSAAAAAAGAGAGAGAGAGAGAGAGAG 1840
Qy 703 GVA-APPGVGLSFFPFGAGCAGKAG 727
Db 1841 GTWGGGGYSDSAAAAAAG 1866

RESULT 13
D70807
Hypothetical glycine-rich protein RV3514 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70807
R:Conor, R.; Brasch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, N.; Cole, S.T.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajendran, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MIM:98295987
A:Accession: D70807
A:Status: Preliminary; nucleic acid sequence not shown; translation: not shown
A:Residues: 11483 <COL>
A:Cross-references: GB:AL02022; GB:AL023456; NID:93261554; PID:GB:AL17751.1; PID:9292
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV3514
C:Superfamily: collagen alpha 1(IV) chain

Query Match 16.5%; Score 626; DB 2; Length 1489;
Best Local Similarity 32.7%; Pred. No. 3.8e-22;
Matches 260; Conservative 26; Mismatches 347; Indels 162; Gaps 35;
Qy 2 GYPCALPGV-PGCVFPGAGCGALG--GALGPGGKPLKPV-----PGLAGAGL 50
Db 119 GANNTTPGAGAGAGLLFNGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174
Qy 51 AGLAGFAPVFPGLVPSG-VADAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 109
Db 175 TG-----GAG 218
Qy 110 VPKKVPGLVPSGVPLGAPFP---GVVLPVPTGAGVKNPAGVGVGAGFAGIPGV 166
Db 219 -----GVGAG 264
Qy 167 GPGGPGVPLPGLPIKAPKAGVGLPPTTKLPGYGGVAGAGAGAGAGAGAGAGAGAGAG 224

Dd	11	GAPOHIGGAGGAGTGLVGGAGGAGGDSNAVHYIGGAGGAGGALLFGAGGAGGAGGSGS	70
QY	53	LGAPVPTFFCALPVGVDADAAAYAAKANG-----AGLGGVGVGGVLSAGAV	102
Dd	71	SGA-----AGRGAGAGGCLFASGSGSGFGFASCTGCGAGAGTGGAGGLFASGG-	120
QY	103	VFOQGAGTKCPVPCVGLPSPVPGCVLGPAPPPGVGLVPPVPTGATKAPGVGCA--F	160
Dd	121	VSGTGGGAGSGTGTGVGSGTG--GAGLFPASGGAGAG-----GSGCTGAGGTVGGAGL	172
QY	161	AGITGTVGPEGSP-----QPCVPLGYPIKAPLPGVGLPPTTKLPYGVGGVAGA	212
Dd	173	FGAGGAGGLGAGCQNTGCGHGGAGGSGAGLLALGCGAGGAGAAT-----GTGAGGA	225
QY	213	AGKAGTPTGT--GVGPQAAAN-----AAKAAKATGAGACVLPFG-----	250
Dd	226	GGAGLGLFGSGGSGGAGTFTGTONSGGAGGAGKAGLFGSGGSGGSGGAGGFGA	285
QY	251	-VWGAGVCPVGPATPGTGTGAGTGPAAARAAAARAAKAKYGAACLVPGCGFGPSTVG	309
Dd	286	STGGAGGAGSGAGLIGNGGSGGTSVATGAGNGCAGGAGGAGLIGNGGSGSGMG	345
QY	310	-VFGAGVPGVGVGAGTVPVFGAGTIPGAPVPGVSPFAAAKAAKAYGARCPVGVGI	368
Dd	346	DAPG-----GTCVGGTIGBLLL-----GLDANNAFATNPLTAQQRALAAVA--	396
QY	369	PTYGAGGFGFGVGVGTPGVAGVPSVGVPGVGVGVGVGVGVGVGVGVGVGVGVGV	428
Dd	397	PLIGNGNGAP-----SGRFGHG-----GMLPGGGTGGSGGAGGDDGAGLIFGA	447
QY	429	GTPPAAARAAKAAKAGFLGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV	486
Dd	448	GCAGGAG-----GAVTGTGATGSGSGGAGGALLFGAGGAGGAGSGTGGPAGG	497
QY	487	VGVVAPGIFGSGV-----AAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA	538
Dd	498	AG-----GPGGAGGIFNGSGGAGGSGVSGGAGGEGAG--GAG--GLFPAGGAGGAG	549
QY	539	GAGVGVGLGVGAGVGV--FGAGADGGVHRHSVPELRQDPSSSHLPSTSPRVPQALAA	597
Dd	550	GNNYGAGGAGGV	587
QY	598	KAAKYGAAPVGLSGALGCVGIPGVGVGAPPAARAAKAAKAAKAAKAAKAAKAAKAA	657
Dd	598	GNA--GLLAPLAGGAGSGSGGQPTGAG--GFGGDAGS-----LVGSGGVGCA	633
QY	658	GVGSL--GVPGV-----GGLGITPPAAAKAAKYGAGLGVGVGAGGQPLGGY	704
Dd	634	GGFELITGGPAGAGDAGLFGSGGAG--AGGSGFTDZGAGGAGGKAGLIGNGNGRA	691
QY	705	AAKFGTGLSPFFGGAQLOKAG	727
Dd	682	GGAGGNGGDDGFGPQARNFELNG	714
Status: nucleic acid sequences not shown; translation not shown			
Result 15			
CGH48			
collagen alpha 1(IV) chain precursor - human			
NAlternate names: procollagen alpha 1(IV) chain			
CSpecies: Homo sapiens (man)			
CDate: 28-May-1986 #sequence_revision 31-Dec-1999 #name_change 07-Dec-1999			
CAccession: S16876; A32117; S02738; S00048; S25826; A33115; S00207; S39614; A02863;			
R.Sjölinen, R.; Huotari, M.; Ganguly, A.; Prockop, D.J.; Tryggvason, K.			
R.Jölin, Chem. 264, 13565-13571, 1989			
A.Title: Structural organization of the gene for the alpha-1 chain of human type IV c			
A.Accession: S16876			
A.Status: nucleic acid sequences not shown; translation not shown			
A.Residues: 1-1669 <SO1>			
A.Cross-references: EMBL:J04217; GE:J05039; NID:g180800; PIDN:RAA53098.1; PID:g180803.			
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1988			
R.Sjölinen, R.; Huotari, M.; Hostikka, S.L.; Prockop, D.J.; Tryggvason, K.			

RESULTS 15

Cont448

Valentine alpha 1(IV) chain precursor - human

Valentine alpha 1(IV) chain

C.Species: human

C.Name: procollagen alpha 1(IV) chain

C.Species: Homo sapiens (man)

C.Date: 26-Feb-1999

C.Reference: #1569, #1569 revision 31-Dec-1992 #test change 07-Dec-1999

C.Accession: S16876; S16877; S16878; S16879; S16880; S16881; S16882; S16883; S16884; S16885; S16886; S16887; S16888; S16889; S16890; S16891; S16892; S16893; S16894; S16895; S16896; S16897; S16898; S16899; S16900; S16901; S16902; S16903; S16904; S16905; S16906; S16907; S16908; S16909; S16910; S16911; S16912; S16913; S16914; S16915; S16916; S16917; S16918; S16919; S16920; S16921; S16922; S16923; S16924; S16925; S16926; S16927; S16928; S16929; S16930; S16931; S16932; S16933; S16934; S16935; S16936; S16937; S16938; S16939; S16940; S16941; S16942; S16943; S16944; S16945; S16946; S16947; S16948; S16949; S16950; S16951; S16952; S16953; S16954; S16955; S16956; S16957; S16958; S16959; S16960; S16961; S16962; S16963; S16964; S16965; S16966; S16967; S16968; S16969; S16970; S16971; S16972; S16973; S16974; S16975; S16976; S16977; S16978; S16979; S16980; S16981; S16982; S16983; S16984; S16985; S16986; S16987; S16988; S16989; S16990; S16991; S16992; S16993; S16994; S16995; S16996; S16997; S16998; S16999; S17000; S17001; S17002; S17003; S17004; S17005; S17006; S17007; S17008; S17009; S17010; S17011; S17012; S17013; S17014; S17015; S17016; S17017; S17018; S17019; S17020; S17021; S17022; S17023; S17024; S17025; S17026; S17027; S17028; S17029; S17030; S17031; S17032; S17033; S17034; S17035; S17036; S17037; S17038; S17039; S17040; S17041; S17042; S17043; S17044; S17045; S17046; S17047; S17048; S17049; S17050; S17051; S17052; S17053; S17054; S17055; S17056; S17057; S17058; S17059; S17060; S17061; S17062; S17063; S17064; S17065; S17066; S17067; S17068; S17069; S17070; S17071; S17072; S17073; S17074; S17075; S17076; S17077; S17078; S17079; S17080; S17081; S17082; S17083; S17084; S17085; S17086; S17087; S17088; S17089; S17090; S17091; S17092; S17093; S17094; S17095; S17096; S17097; S17098; S17099; S17100; S17101; S17102; S17103; S17104; S17105; S17106; S17107; S17108; S17109; S17110; S17111; S17112; S17113; S17114; S17115; S17116; S17117; S17118; S17119; S17120; S17121; S17122; S17123; S17124; S17125; S17126; S17127; S17128; S17129; S17130; S17131; S17132; S17133; S17134; S17135; S17136; S17137; S17138; S17139; S17140; S17141; S17142; S17143; S17144; S17145; S17146; S17147; S17148; S17149; S17150; S17151; S17152; S17153; S17154; S17155; S17156; S17157; S17158; S17159; S17160; S17161; S17162; S17163; S17164; S17165; S17166; S17167; S17168; S17169; S17170; S17171; S17172; S17173; S17174; S17175; S17176; S17177; S17178; S17179; S17180; S17181; S17182; S17183; S17184; S17185; S17186; S17187; S17188; S17189; S17190; S17191; S17192; S17193; S17194; S17195; S17196; S17197; S17198; S17199; S17200; S17201; S17202; S17203; S17204; S17205; S17206; S17207; S17208; S17209; S17210; S17211; S17212; S17213; S17214; S17215; S17216; S17217; S17218; S17219; S17220; S17221; S17222; S17223; S17224; S17225; S17226; S17227; S17228; S17229; S17230; S17231; S17232; S17233; S17234; S17235; S17236; S17237; S17238; S17239; S17240; S17241; S17242; S17243; S17244; S17245; S17246; S17247; S17248; S17249; S17250; S17251; S17252; S17253; S17254; S17255; S17256; S17257; S17258; S17259; S17260; S17261; S17262; S17263; S17264; S17265; S17266; S17267; S17268; S17269; S17270; S17271; S17272; S17273; S17274; S17275; S17276; S17277; S17278; S17279; S17280; S17281; S17282; S17283; S17284; S17285; S17286; S17287; S17288; S17289; S17290; S17291; S17292; S17293; S17294; S17295; S17296; S17297; S17298; S17299; S17300; S17301; S17302; S17303; S17304; S17305; S17306; S17307; S17308; S17309; S17310; S17311; S17312; S17313; S17314; S17315; S17316; S17317; S17318; S17319; S17320; S17321; S17322; S17323; S17324; S17325; S17326; S17327; S17328; S17329; S17330; S17331; S17332; S17333; S17334; S17335; S17336; S17337; S17338; S17339; S17340; S17341; S17342; S17343; S17344; S17345; S17346; S17347; S17348; S17349; S17350; S17351; S17352; S17353; S17354; S17355; S17356; S17357; S17358; S17359; S17360; S17361; S17362; S17363; S17364; S17365; S17366; S17367; S17368; S17369; S17370; S17371; S17372; S17373; S17374; S17375; S17376; S17377; S17378; S17379; S17380; S17381; S17382; S17383; S17384; S17385; S17386; S17387; S17388; S17389; S17390; S17391; S17392; S17393; S17394; S17395; S17396; S17397; S17398; S17399; S17400; S17401; S17402; S17403; S17404; S17405; S17406; S17407; S17408; S17409; S17410; S17411; S17412; S17413; S17414; S17415; S17416; S17417; S17418; S17419; S17420; S17421; S17422; S17423; S17424; S17425; S17426; S17427; S17428; S17429; S17430; S17431; S17432; S17433; S17434; S17435; S17436; S17437; S17438; S17439

[illegible]

F:1392,1395,1398,1404/Modified site: 4-hydroxyproline (Pro) #status experimental
F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted
F:1505-1511,1616-1622/Disulfide bonds: #status predicted
F:1570-1602,1604-1605/Disulfide bonds: (or 1570-1605, 1604-1602) #status predicted

Query Match 16.1%; Score 609.5; DB 1; Length 1669;
Best Local Similarity 28.8%; Pred. No. 2.4e-21;
Matches 265; Conservative 63; Mismatches 286; Indels 295; Gaps 54;

```
QY 2 GYEGAI---PGVPGSVYFPGA-----GLGALG-----GGALG 31
DB 574 GSPGVGKXGKSGPGVPGSGRDPFGPGFAGFAGKQAGQAGPGPGSGPLG 633
QY 32 PGKPK-K--FVGLAGAGAGAGAPAVT---FGLAVGVNDAAAKAKAKAG 85
DB 634 PKSGFKIVLPSPGAGL-PGSPFGPGGGRGFS--FGRPLFSEKAGVGPGLG 690
QY 86 LGVGVGSLGVSAGVPGAGVPG--KVPGLPSV--YRGLV 129
DB 691 FPGPGKVGDLPGDMGP-FGPGRGNGLGPNVGQKGPGLGLKGLF-L 747
130 PGARFPG-----VGLVGVFGAGY-----KPRAPGVGNAF--GIPGVGP 169
DB 748 PG--IFGTGEXGSLGV-PVPGHGAIGPGLQIRGEFPGPLPSGVSPGICPP 804
QY 170 GGFPGVPLGPIKAFKLGGLTLYTKLFYGVGVGAGAKAGAPGTGVPDAA 229
DB 805 GARGP-----PGQSPGLS-----GPPILGKGFPPGLNRP--- 841
QY 230 AAAAAAAGAGAGVLPV--GAGVGVPG--ALPILGAG--VCTPAAA 279
DB 842 -----KQKGAQ--LPGITGSLGLPQQAQAPGPPGSKGEMGAT----- 886
QY 280 AAAAAAAGAGAGVLPVPGVPGVGA-----GVPGV---GVPG----- 322
DB 887 -----PGQPG--SPGPGAPGLGKGDHGPSSGPRGDLGKRG 927
QY 323 -AGTPVPGA-----GIPGANVGVSPERAAKA 350
DB 928 DVLGPKSGKXNDGSKQKQDQKQIGFGEKSGRDPG--TPGFSQS--Q 982
QY 351 AKAAGYAGRGVGVGIPYGVAGGPG---FGVGVGIPGVAGVPGVGVGVS 405
DB 983 AGQPGQPGKDPGIGSTP-----GAPGLGPKSVGAGLPGTEGKGP---GIRPGQ 1035
QY 406 VFGVGSPEAQAAX--AAKYGVTPAAAAAKAAQAGLIVG-----VGVP 456
DB 1036 SPGL--FGDKAGKGGAGPGGIGPLRGEKGDQAGFPGSGEKESIGI-PG 1091
457 VGVAGVGPAGC-VGL-----APGVVAGVGVAGVAGVAG-IGGCVAA 500
DB 1092 MFGSPGLGSPSVGVPGSGPLGKDKLPLGLGVKENGILPCTGTGTPAGQG 1151
QY 501 AKAAGYAAKALRAAGL-GAGILGLVG-----VGVRLGVAGVPG----- 544
DB 1152 EFGSDG--FGSAGEKGPGLGPGFPFGAKGDKSGKNGFPLGAGSGIPGSGEQ 1210
QY 545 -LVGA-----GVPGFAGAGVRSRLREDDPSQHLSTPSPRPVPG--ALAA 596
DB 1211 FVGPGPGQPGPLPGSGHATGPGKGRGPGGQGLPG---LPGMPPGLGIDYK 1266
QY 597 AKAKYAAAGVYGLGLGLGVGVGAGPAAAAAKAAKAAQGLVAGLGG 656
DB 1267 DXGPGWAGVPGKGPFGQNGP--IGSGPLTGSNGMPPGVP-GFQPGKLP 1323
QY 657 L-GV---GGLVGVGVGLGIP-----PAAAAAKYGAAGLGVGGAGQF 699
DB 1324 LQKIGDQDQGVGAGVPGFPGPGPDYIKGPFGLPGGLGQGLGKPGQ 1383
QY 700 PLGGVAAKPGGLSIFPG 718
DB 1384 GTGLGIPG---PPGPG 1399
```

Search completed: April 24, 2001, 16:41:50
Job time: 455 sec

DB	578	VSPAAAARAAK7GAAGSLAGVLGAGQFFTIGGAGGGLVGKPKPFGGALGALGFFGGNC	737
QY	722	LNKAGGNKKK 731	
		1147111111	
DB	738	LNKAGGNKKK 747	
RESULT 3			
ID	ELG_MOUSE	STANDARD:	PRT: 860 AA.
AC	F54320:		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-OCT-1996 (Rel. 34, Last annotation update)		
DE	ELASTIN PRECURSOR (TROPOLASTIN).		
OS	ELN.		
OS	Mus musculus (Mouse).		

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N. A.

```
CC STRAIN-BAL/C; TISSUE-Lung;
```

```
RX MEDLINE=95130659; PubMed=7829060;
```

```
RA Wyder K.S.; Secher J.L., Boyd C.D., Passmore H.C.:
```

```
RT Use of an action polymerase to localise the tropoelastin gene to
```

```
RM chromosome 7 in a region of linkage conservation with Human
```

```
RL Chromosome 7 " " "
```

```
RC Genomics 23:125-131(1994).
```

```
-! FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
```

```
CC NUCLEON LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
```

```
-! SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
```

```
CC INTO AN EXTENSIBLE 3D NETWORK.
```

```
-! SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
```

```
CC -! PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYLS.
```

```
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
```

```
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by for commercial
```

```
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to licens@isb-sib.ch).
```

```
CC EMBL; D040210; AA80155.1 ;
```

```
DR HSPP; U040021; IWA.
```

```
DR Structural Protein; Repeat; Signal; Connective tissue.
```

```
KW POTENTIAL.
```

```
FT CHAR. 26 82
```

```
FT SEQUENCE 860 AA; 71955 WA; QCGEBSAEIEDYFI CRC64;
```

```
Query Match 63.1%; Score 2387; DB : Length 860;  
Best Local Similarity 61.1%; Pred. No. 1.le-87;  
Matches 546; Conservative 25; Mismatches 96; Indels 226; Gaps 34;
```

```
QY      1 GVTPGARFGGVGGVGYPCAGLAGL--GGSGLGKGCKPLKP-----VPGLAGAG 48  
Db       32 GAAPFSGARPGPVNYPVGALIVGSLGGEGGKPKRPNPKNGLTGTFGAQGVLGGAG 91  
QY     49 LGASLGARGANVFPR---CALPYGTGGTAABAAAYK-AAKAAGAAGAGGVGVSY    92  
Db        92 FCAGILGRFAPICTFFNGALVFGSAACAANAATAAKAGAGLGGVGGVGGVGPVGGV 151  
QY           -----GGLYGSAVVYFPFGAGV----KKQKPFGLPGVY 124  
Db          152 GVGSPFGGVGGVGGVGGVGGISGLGSLATGVAITPVQVGAAGIAGACKTKYPTFLPGVY 211  
QY         125 PGVLIP---GARFGGVGVLPSTVTIGAGVVKPAKRGVGNAGIFGFVGFGFGLPYTI 182  
Db          212 FGAVLTGTCAKRFPGVGLPSYTVTIGVVMARKGGAGAGSSLVGFGVGGQQGVGLPYTI 271
```


ID	CALA_HUMAN	STANDARD:	FRT; 1659 AA.
P02482	1996 (Rel. 01. Created)		
01-SEP-1996	(Rel. 33. Sequence update)		
01-SEP-2000	(Rel. 40. Last annotation update)		
01-OCT-2000	(Rel. 40. Last annotation update)		
COLLAGEN ALPHA 1(IV) CHAIN PRECURSOR.			
COLAAL			
Homo sapiens (Human).			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
NCBI_TaxID=9606;			
[1]	SEQUENCE FROM N.A.		
MDLIN6-859340433;	PubMed=2701944;		
Soininen R., Huotari M., Ganguly A., Prockop D.J., Tryggyvason K.;			
Structural organization of the gene for the alpha 1 chain of human			
Type IV collagen.			
[2] Biol. Chem. 264:13565-13571 (1989).			
[3]	SEQUENCE OF 46-1257 FROM N.A.		
TISSEB-Placenta;			
MDLIN6-86083584;	PubMed=3691802;		
Soininen R., Haka-Risku T., Prockop D.J., Tryggyvason K.;			
"Complete primary structure of the alpha 1-chain of human basement			
membrane (type IV) collagen."			
FEBS Lett. 225:188-194 (1987).			
[4]	SEQUENCE OF 1-943 FROM N.A.		
MDLIN6-8609471;	PubMed=3311751;		
Dreitzmann R., Kuehn K., Dieringer H., Babel W., Glanville R.W.,			
Prockop D.J., Kuehn K.;			
"Completion of the amino acid sequence of the alpha 1 chain of human			
basement membrane collagen (type IV) reveals 21 non-triplet			
interruptions located within the collagenuous domain."			
Eur. J. Biochem. 168:529-536 (1987).			
[5]	SEQUENCE OF 28-243.		
MDLIN6-86004708;	PubMed=4043082;		
Glanville R.W., Qian K.Q., Steeb B., Risteli J., Kuehn K.;			
"The amino acid sequence of the N-terminal triple helix-forming			
region (7S collagen) of the alpha 1 (IV) chain of human basement			
membrane collagen."			
Eur. J. Biochem. 152:213-219 (1985).			
[6]	SEQUENCE OF 534-1447.		
MDLIN6-85003629;	PubMed=6434307;		
Babel W., Glanville R.W.;			
"Structure of human-basement-membrane (type IV) collagen. Complete			
amino-acid sequence of a 914-residue-long peptic fragment from the			
alpha 1(IV) chain." Eur. J. Biochem. 143:545-556 (1984).			
[7]	SEQUENCE OF 1256-1669 FROM N.A.		
MDLIN6-85207819;	PubMed=2581969;		
Pitkilahti T., Tryggyvason K., Myers J.C., Kurkinen M., Lebo R.;			
Cheng M.-C., Prockop D.J., Boyd C.D.;			
"cDNA clones coding for the pro-alpha1(IV) chain of human type IV			
procollagen reveal an unusual homology of amino acid sequences in two			
halves of the carboxyl-terminal domain."			
Eur. J. Biochem. 260:7661-7687 (1985).			
[8]	SEQUENCE OF 1255-1669 FROM N.A.		
MDLIN6-85714528;	PubMed=2581969;		
Prinkner U.M., Rodan G., Foidl H.R., Wang S.-Y., Rosenblom J.;			
Kerfalis W.A., Myers J.C.			
"Restricted homology between human alpha 1 type IV and other			
procollagen chains."			
Proc. Natl. Acad. Sci. U.S.A. 82:3649-3653 (1985).			

RP
SEQUENCE OF 1-28 FROM N.A.
RX MEDLINE=89034231; PubMed=3182844;
RA Soininen R., Huotari M., Hostikka
S.I., Prockop D.J., Triggavason K.;

Db 1015 GQPLGIFPG---LKGTIG-DMGFFGPGVGEPRG---PSSVPGQPG 1054
 QY 548 GAGVPG-FGAGADEGVRASLSPEL--RQDPSSSOHLPTSPSSPVRGA-----LAA 596
 Db 1055 SPFLGPGKMGKDGIGSLGSLGPGKGEFG-----LPGYGNPGIKGSGVGPGLGPG 1110
 QY 597 ARAAKYGAAPGVGLGALGALGVGI-----PGTVGAGPAAAAAAAKAA 644
 Db 1111 TFGKQPGPGLFFGPGFGKFGISPGNGNPGLPSPGPGVGGGHGPGQPGKPKPG 1170
 QY 645 PGLGVGARG-LGSLGVGLGVGVLGSLGIPPAANAANKYAGAGLVGAGQFPLGS 703
 Db 1171 OGSLPGKAGKEGPGQFGNGPGLGEL-----SQKGGGLGPGPNRG---LPG 1220
 QY 704 VAARPGF-GLSPIFG 718
 Db 1221 PKGPGFPG-----PFG 1232
 Db 1222 PKGPGFPG-----PFG 1232
 RESULT 13
 CHAIN_MOUSE STANDARD: PRT: 1669 AA.
 AC 021463;1986 (Rel. 01. Created)
 DT 01-FEB-1991 (Rel. 17. Last sequence update)
 DT 15-JUN-1999 (Rel. 38. Last annotation update)
 DE COLLAGEN ALPHA 1(IV) CHAIN PRECURSOR.
 GN COLA1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI-TaxID=10090;
 RN [1] JENANCE FROM N.A.
 RN [2] MEDLINE=89197932; PubMed=2703490;
 RN [3] Muthukumar G., Blumberg B., Kurkinen M.;
 RN [4] "The complete primary structure for the alpha 1-chain of mouse
 RN [5] collagen IV. Differential evolution of collagen IV domains.";
 RN [6] J. Biol. Chem. 264:6310-6317(1989).
 RN [7] SEQUENCE OF 1-1154 FROM N.A.
 RN [8] MEDLINE=88112221; PubMed=3338568;
 RN [9] Wood L., Theriault N., Vogeli G.;
 RN [10] "cDNA clones completing the nucleotide and derived amino acid
 RN [11] sequence of the alpha 1 chain of basement membrane (type IV) collagen
 RN [12] from mouse.";
 RN [13] J. Biol. Chem. 264:227-234(1989).
 RN [14] SEQUENCE OF 1149-1424 FROM N.A.
 RN [15] MEDLINE=86301886; PubMed=3756692;
 RN [16] Nath P., Laurent M., Horn E., Schel M.E., Zon G., Vogeli G.;
 RN [17] "Isolation of an alpha 1 type-IV collagen cDNA clone using a
 RN [18] synthetic oligodeoxynucleotide.";
 RN [19] J. Biol. Chem. 263:301-304(1988).
 RN [20] SEQUENCE OF 1276-1669 FROM N.A.
 RN [21] MEDLINE=85127033; PubMed=2574961;
 RN [22] Verbaeten L., Laurent M., Schwarz U., Sakurai Y., Yamada Y.,
 RN [23] "cDNA clones completing the nucleotide and derived amino acid
 RN [24] sequence of the alpha 1 chain of basement membrane (type IV) collagen
 RN [25] from mouse.";
 RN [26] J. Biol. Chem. 263:301-304(1988).
 RN [27] SEQUENCE OF 1441-1669 FROM N.A.
 RN [28] MEDLINE=87250460; PubMed=3597383;
 RN [29] Kurkinen M., Condon M.R., Blumberg B., Barlow D., Quinones S.,
 RN [30] Saus J., Pihlajaniemi T.;
 RN [31] "Extensive homology between the carboxyl-terminal peptides of mouse
 RN [32] alpha 1(IV) and alpha 2(IV) collagen.";
 RN [33] J. Biol. Chem. 263:5496-5499(1988).
 RN [34] PARTIAL SEQUENCE FROM N.A.

RX MEDLINE=86196099; PubMed=3009468;
 RA Sakurai Y., Sullivan M., Yamada Y.;
 RT "Alpha 1 type IV collagen gene evolved differently from fibrillar
 RT collagen genes.";
 RL J. Biol. Chem. 261:6654-6657(1986).
 RN [7]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=89065738; PubMed=3198626;
 RA Kayes P., Wood L., Theriault N., Kurkinen M., Vogeli G.;
 RT "Head-to-head arrangement of murine type IV collagen genes.";
 RL J. Biol. Chem. 263:19274-19277(1988).
 RN [8]
 RP SEQUENCE OF 1-28 FROM N.A.
 RX MEDLINE=89071759; PubMed=3200851;
 RA Burdello P.D., Martin G.R., Yamada Y.;
 RT "Alpha 1(IV) and alpha 2(IV) collagen genes are regulated by a
 RT bidirectional promoter and a shared enhancer.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9579-9582(1988).
 RN [9]
 RP SEQUENCE OF 1-129 FROM N.A.
 RX MEDLINE=88243724; PubMed=3379041;
 RA Killen P.B., Burdello P., Sakurai Y., Yamada Y.;
 RT "Structure of the amino-terminal portion of the murine alpha 1(IV)
 RT collagen chain and the corresponding region of the gene.";
 RL J. Biol. Chem. 263:8787-8791(1988).
 CC [1] FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
 CC GLomerular BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE/
 CC MESHWORK TOGETHER WITH LAMININS, PROTEGLYCANS AND ENACTIN/
 CC NIDOGN.
 CC [2] SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
 CC ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
 CC WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
 CC [3] DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
 CC DOMAIN (NCL) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
 CC CHAIN BY HYDROXYLATIONS, AND A SHORT N-TERMINAL
 CC CAUSE RESIDUES IN THE CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
 CC TRIPLE-HELICAL) IS DOMAIN.
 CC [4] PM: PROLINES AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC [5] PM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
 CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
 CC THESE, LOCATED IN THE NCL DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
 CC IV COLLAGENS.
 CC [6] THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use, but the EMBL outstation is not responsible for any use of the
 CC modified and this statement is not removed. Please use the EMBL
 CC outstation to request a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC [7] EMBL: J03758; AA37439.1; -
 CC EMBL: M23333; AA51625.1; -
 CC EMBL: J04694; AA50292.1; -
 CC EMBL: X06777; CA29946.1; -
 CC EMBL: X02201; CA26132.1; -
 CC EMBL: M15832; AA37340.1; -
 CC EMBL: M1042; AA37342.1; -
 CC EMBL: M13027; AA37342.1; -
 CC EMBL: M13025; -; NOT ANNOTATED_CDS.
 CC EMBL: M13026; AA37344.1; -
 CC EMBL: M13027; AA37345.1; -
 CC EMBL: M13043; AA37346.1; -
 CC EMBL: J04448; AA37437.1; -
 CC PIR: A33525; CQMS48.
 CC MGD: MG1:88454; Col4a1.
 CC InterPro: IPR000087; -
 CC InterPro: IPR001442; -
 CC Pfam: PF04413; C4; 2.
 CC Pfam: PF03391; Collagen; 21.
 CC Extracellular matrix; Connective tissue; Basement membrane;

RP MEDLINE=89030591; Pubmed=2745420;
RA Marini J.C., Grange D.X., Gottesman G.S., Lewis M.B., Koepfli D.A.;
RX "Osteogenesis imperfecta type IV: detection of a point mutation in
RT the alpha 1(I) collagen allele (COL1A1) by RFLPA hybrid analysis.";
RL J. Mol. Chem. 264:11895-11901(1989). [21]
RP VARIANTS OI-II 24A-11106; VAL-11501; ARG-11154 AND VAL-11184.
RX MEDLINE=89380165; Pubmed=2777764;
RA Lemande S.R., Dahl H.-B.M., Cole W.G., Bateman J.F.;
RX "Characterization of point mutations in the collagen COL1A1 and
RT COL1A2 genes causing lethal perinatal osteogenesis imperfecta.";
RL J. Biol. Chem. 264:15809-15812(1989). [22]
RP VARIANTS OI-III 25A-11065; VAL-11065 AND VAL-11066.
RX MEDLINE=90062068; Pubmed=2511192;
RA Schalk T., Constantinou C.D., Kella K., Nielsen K.B., Prockop D.J.;
RX "Structural variants of osteogenesis imperfecta: molecular genetic
RT variant of osteogenesis imperfecta naturally stabilizes the triple
EL helix of type I procollagen. The effects of glycine substitutions on
thermal stability are either position or amino acid specific.";
RL J. Biol. Chem. 264:15694-15699(1989). [23]
RP VARIANT OI-IV SER-1022.
RX MEDLINE=90081055; Pubmed=2913053;
RA Constantinou C.D., Nielsen K.B., Prockop D.J.;
RX "A lethal variant of osteogenesis imperfecta has a single base
RT mutation that substitutes cysteine for glycine 904 of the alpha 1(I)
EL chain of type I procollagen. The asymptomatic mother has an
unstable procollagen.";
RL J. Clin. Invest. 83:574-584(1989). [24]
RP VARIANT OI-VS-272; CYS-704 AND CYS-896.
RX MEDLINE=90009313; Pubmed=2794057;
RA Starnan B.J., Eyre D., Charbonneau H., Harrylock M., Wells M.A.,
RX Weiss L., Graham J.N., Evers P.H.;
RX "Osteogenesis imperfecta. The position of substitution for glycine by
RT cysteine in the triple helical domain of the pro alpha 1(I) chains of
EL type I collagen determines the clinical phenotype.";
RL J. Clin. Invest. 84:1206-1214(1989). [25]
RP VARIANT OI-II CYS-422.
Query Match 15.3%; Score 580; DB 1; Length 1464;
Best Local Similarity 33.7%; Pred. N. 1.1e-16;
Matches 268; Conservative 3%; Mismatches 342; Indels 154; Gaps

QY 2 GVPGAGTGGVGVGFTFG-AGL-GALG-GGALGGKGKPLP-----VPGLAGAGAGAG 54
Db 254 GLPTF-AGLFGKHGRGFSGLDGAAGDPAGFGPSGFSGPMGAPMGPPGPRG 311
QY 55 APFAVTFFGALLRVSIV-----ADAAAKYKAAGAAGLVGVGLCGLSAGNVPWSPG 109
Db 312 R-----DGAPGAGAGNGCATGAGPGPTGPAGPGFPFGVGANGEGPCPGPSBS 365
QY 110 VPKKKVPG-VOLPVNVGVLPGARFEVGVLGVPTGAGVKFKAPGVGAAGLPGV-G 167
Db 366 ----PGVVRCEPFGP-PAGAGAPAGPAGDGOFGKANG-----APCIAGA-FGPGGAG 416
QY 168 FFGGPFQFGVLGYPIAFKLFGYGLFTTKLPLGVGVGVGAGAAKAGATFGTGV-GP 226
Db 417 FGSGQFGGGPGF-----FK-GNSGSEP-----GAPGSKGDGTAKGSPGVGQP 459
QY 227 QAAAANAANKANAGAGAGVLPV-----GGAGVGVPGR-IPTIGLIAGR-GPAAA 279
Db 279 QAAAANAANKANAGAGAGVLPV-----GGAGVGVPGR-IPTIGLIAGR-GPAAA 279
QY 460 PGPAGECKNGKRGVFPG-LPGFPRGPDGSGKFFDQADVAGPAGPAGGAGPAG 518
Db 518 PGPAGECKNGKRGVFPG-LPGFPRGPDGSGKFFDQADVAGPAGPAGGAGPAG 518
QY 280 AAAAAAANKYIS-----AAGAVL-PGGGFGFGVGVPGA-----GVPG-VGVPG- -A 323
Db 323 AAAAAAANKYIS-----AAGAVL-PGGGFGFGVGVPGA-----GVPG-VGVPG- -A 323
QY 519 PKGSFPEAGRGEAGLPGAMLTGSPSGF-PDKTGTPPGAGCQHPGPPGPGAGQA 577
Db 577 PKGSFPEAGRGEAGLPGAMLTGSPSGF-PDKTGTPPGAGCQHPGPPGPGAGQA 577
QY 324 GIPIVVFG-----AGTIG-RHVGVGPSAAKAAAKAYKAPGVGVGIFITTVGAGCF 378
Db 378 GIPIVVFG-----AGTIG-RHVGVGPSAAKAAAKAYKAPGVGVGIFITTVGAGCF 378

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:36:20 ; Search time 62.39 seconds
(without alignments)
61.891 Million cell updates/sec

Title: US-09-340-736-2

Perfect score: 1007

Sequence: 1 PFPGVGVGPGVGVGV.....GVAFGVGVAPGVGVAPGVAP 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched:

185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Vatch 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

2: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

3: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

4: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

5: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

6: /cgn2.5/ptodata/2/iaa/5A.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1037	100.0	201	2	US-08-911-364-2
2	700	68.8	731	2	US-08-911-364-1
3	662.5	65.1	792	2	US-08-678-039A-40
4	401	39.4	988	1	US-08-212-237-5
5	397.5	39.1	972	1	US-08-212-237-7
6	397.5	39.1	972	1	US-08-212-237-7
7	397.5	39.1	972	1	US-08-212-237-7
8	392	38.5	1056	1	US-08-212-237-6
9	391.5	38.5	1056	1	US-08-212-237-6
10	391.5	38.5	1056	1	US-08-212-237-6
11	390	38.3	1024	1	US-08-212-237-8
12	390	38.3	1024	1	US-08-212-237-8
13	390	38.3	1024	1	US-08-212-237-8
14	382	37.6	936	1	US-08-212-237-4
15	382	37.6	936	1	US-08-212-237-3
16	379.5	37.3	1413	1	US-08-175-155-39
17	379.5	37.3	1413	1	US-08-175-155-39
18	379.5	37.3	1413	1	US-08-175-155-39
19	379.5	37.3	1464	1	US-08-477-509B-74
20	379.5	37.3	1464	1	US-08-477-509B-74
21	376.5	37.0	281	1	US-08-482-085B-74
22	376.5	37.0	281	1	US-08-482-085B-74
23	376.5	37.0	281	1	US-08-482-085B-74
24	376.5	37.0	281	1	US-08-482-085B-74
25	371.5	36.5	2357	1	US-08-477-509B-74
26	371.5	36.5	2357	1	US-08-477-509B-74
27	368.5	36.2	378	2	US-08-707-237A-104

Sequence 26, Appl
Sequence 26, Appl
Sequence 103, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 46, Appl
Sequence 81, Appl
Sequence 52, Appl
Sequence 81, Appl
Sequence 68, Appl
Sequence 15, Appl
Sequence 96, Appl
Sequence 15, Appl
Sequence 12, Appl
Sequence 77, Appl
Sequence 73, Appl
Sequence 78, Appl
Sequence 9, Appl

ALIGNMENTS

RESULT 1
US-08-911-364-2

: Sequence 2, Application US/08511364

: Patent No. 5969106

: GENERAL INFORMATION:

: APPLICANT: ROTHSTEIN, Aser

: APPLICANT: KEELY, Fred W.

: TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN

: TITLE OF INVENTION: ELASTIN AND OTHER FIBROUS PROTEINS

: NUMBER OF SEQUENCES: 8

: CORRESPONDENCE ADDRESS:

: STEVEN ROYCE & LARDNER

: STREET: 3000 K Street, N.W.

: CITY: Washington

: STATE: D.C.

: COUNTRY: U.S.A.

: ZIP: 20007-5109

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patentin Release #1.0, Version #1.30

: CURWARE: Patentin Release #1.0, Version #1.30

: APPLICATION NUMBER: US/08/911.364

: FILING DATE: 07-AUG-1997

: CLASSIFICATION: 530

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: US 60/023,552

: FILING DATE: 07-AUG-1996

: ATTORNEY/AGENT INFORMATION:

: NAME: Bent, Stephen A.

: REGISTRATION NUMBER: 29,788

: REFERENCE/DOCKET NUMBER: 041082/0104

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: (202) 674-5100

: TELEFAX: (202) 674-5139

: INFORMATION FOR SEQ ID NO: 2:

: SEQUENCE CHARACTERISTICS:

: TYPE: amino acid

: LENGTH: 201 amino acids

: STRANDEDNESS:

: TOPOLOGY: linear

: MOLECULE TYPE: peptide

: US-08-911-364-2

Query Watch 100.0%; Score 1017; DB 2; Length 201;
Best Local Similarity 100.0%; Prod No. 3 9e-77;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FFGFVGVGIGPVAGVPGVGVSPGVSIPFAQAAAAAKAAKYGVGTAAAAAK 60
|||||
D8 378 FFGFVGVGIGPVAGVPSVSGVGPVGSGVPVGISSIPFAQAAAAAKAAKYGVGTAAAAAK 437

[illegible]

REGISTRATION NUMBER: 20, 015
 REFERENCE/DOCKETS NUMBER: FP-5847-1-PC,BIR
 TELEPHONE: 415-781-1989
 TELEPHONE: 415-781-1989
 TELEFAX: 415-398-3249
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 972 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Protein
 PRT-0595-02772-7

Query Watch 39.1%; Score 397.5; DB 5; Length 972;
 Best Local Similarity 50.0%; Pred. No. 1.4e-25;
 Matches 126; Conservative 8; Mismatches 53; Indels 65; Gaps

2 PEGVG-GVG---GIPGVAGVPGVG---GVPGVG---GVPGVISPERRAARAAKAAK 48
 17 PEGVGPGVGPGVGPGVG-GVPGVGPGVGPGVGPGVGPGVGPGVGAGSGAGSGAG 75
 49 VGVGTPPAAARAAK-----AKAAQGLPLVGVGVGPGVGPGVGPGVGPGVGAG 97
 76 AGSGAGAGSGAGSGAGSGAGSGAGSGAGSGVPGVG-PGVGV-PGVGV-PGVGV-PG 131
 98 NGV-----APGVGV-----APAIPPEQAARAAAKAKYGVTPPAAARAAK- 144
 132 NVGPGVPGVGPGVGPGVGPGVGPGVGPGVGPGVGPGVGAGSGAGSGAGSGAG 191
 145 -----AKAAQGLPLVGVGVGPGVGPGVGPGVGPGVGPGVGPGVGAGV 246
 192 SGAGAGSGAGAGSGAGSGAGSGAGSGAGSGAGSGVPGVG-PGVGV-PGVGV-PGV 297
 189 VAPGVGVAPAG 200
 247 V-PGVGV-PGVG 256

RESULT 8
 US-08-212-237-6
 : Sequence 6, Application
 : Patent No. 5606019
 : GENERAL INFORMATION:
 APPLICANT: Cappello, Joseph

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

```
/ LENGTH: 1056 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-08-212-237-6

Query Match          38.5%; Score 392; DB 1; Length 1056;
Best Local Similarity 52.8%; Pred. No. 4,3e-25;
Matches 121; Conservative 7; Mismatches 57; Indels 44; Gaps 18;

OY 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 53
Db 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 60
OY 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 102
Db 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 115
OY 54 PAAAAAKA-----AKAAQFGLVPGVAPGVPGVAPGVPGVAPGVAP 102
Db 61 GAGSGAGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 174
OY 103 GGVVARGVAPAGPPPPQAAAAAKKXVGTAAAAAKA-----AKAAQF 151
Db 116 GGVV-PGVGPGVGGAGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 174
OY 152 GLVPGVGVAGVGVAGVGVAGVGVAGVGVAGVGVAGVGVAGVAPAG 200
Db 175 GSVPGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV 216

RESULT 9
PCT-US95-02772-6
/ Sequence 6, Application PCT/US9502772
/ GENERAL INFORMATION:
/ TITLE OF INVENTION: Synthetic Proteins As Implantables
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
/ STREET: Four Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/ ADDITIONAL FILE: Floppy disk
/ OPERATING SYSTEM: PC DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US95/02772
/ FILING DATE:
/ CLASSIFICATION:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rowland, Bertram I
/ REGISTRATION NUMBER: 20,015
/ REFERENCE/DOCKET NUMBER: FP-38847-1-PC/BIR
/ TELEPHONE: 415-398-3249
/ TELEFAX: 415-398-3249
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1056 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US95-02772-6

Query Match          38.5%; Score 392; DB 5; Length 1056;
Best Local Similarity 52.8%; Pred. No. 4,3e-25;
Matches 121; Conservative 7; Mismatches 57; Indels 44; Gaps 18;
```

```
OY 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 53
Db 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 60
OY 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 102
Db 2 PGFGVGGIGPVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKKXVGT 115
OY 54 PAAAAAKA-----AKAAQFGLVPGVAPGVPGVAPGVPGVAPGVAP 102
Db 61 GAGSGAGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 174
OY 103 GGVVARGVAPAGPPPPQAAAAAKKXVGTAAAAAKA-----AKAAQF 151
Db 116 GGVV-PGVGPGVGGAGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 174
OY 152 GLVPGVGVAGVGVAGVGVAGVGVAGVGVAGVGVAGVGVAGVAPAG 200
Db 175 GSVPGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV 216

RESULT 10
US-08-212-237-8
/ Sequence 8, Application US/08212237
/ GENERAL INFORMATION:
/ TITLE OF INVENTION: Synthetic Proteins As Implantables
/ NUMBER OF SEQUENCES: 9
/ CORRESPONDENCE ADDRESS:
/ ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
/ STREET: Four Embarcadero Center, Suite 3400
/ CITY: San Francisco
/ STATE: CA
/ COUNTRY: U.S.A.
/ ZIP: 94111-4187
/ COMPUTER READABLE FORM:
/ ADDITIONAL FILE: Floppy disk
/ OPERATING SYSTEM: PC DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/212,237
/ FILING DATE: 11-MAR-1994
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Rowland, Bertram I
/ REGISTRATION NUMBER: 20,015
/ REFERENCE/DOCKET NUMBER: A-58847/BIR
/ TELEPHONE: 415-398-3249
/ TELEFAX: 415-398-3249
/ INFORMATION FOR SEQ ID NO: 8:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1024 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
US-08-212-237-8

Query Match          38.5%; Score 391.5; DB 1; Length 1024;
Best Local Similarity 47.0%; Pred. No. 4,6e-25;
Matches 126; Conservative 8; Mismatches 57; Indels 77; Gaps 21;

OY 2 PGFGV-GVG---GIRGVAGVPGVG---GVPGVG---GVPGVGSIPERQAAAAAKK 48
Db 37 PGFGVPGVPGVPGVPGV-GVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 95
OY 49 YGVGTPAAAAAKA-----AKAAQFGLVPGVGV---APGVGV---APGVGV- 88
Db 96 AGSGAGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 155
OY 89 ---APEVL---APGVGV---APGVGV---APGVGV---APGVGV---APGVGV- 129
Db 156 GVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 215
```

QY 130 AKYGVTPAAAKA-----AKAAQGLVPGVAGVAGVAGV 172
 Db 216 AGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 272
 QY 173 PGVGLAGVAGVAGVAGVAGVAGVAGVAGVAGVAGV 200
 Db 273 PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV 296

RESULT 11
 PCT-US95-02772-8
 ; Sequence 4, Application PC/TUS9502772
 ; GENERAL INFORMATION:
 ; APPLICANT: Protein Polymer Technologies, Inc.
 ; TITLE OF INVENTION: Synthetic Proteins As Implantables
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 ; STREET: Four Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/02772
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Berttram I
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: P-58847-1-PC/BIR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-781-1989
 ; TELEFAX: 415-398-3249
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1024 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 PCT-US95-02772-8

Query Match 38.5%; Score 391.5; DB 5; Length 1024;
 Best Local Similarity 47.0%; Pred. No. 4.6e-25;
 Matches 126; Conservative 8; Mismatches 57; Indels 77; Gaps 21;
 QY 2 PGVGV-GVG-----GIPVAGVPGV-----GVPGVGSPERQAAAAAKA 48
 Db 37 PGVGVPGVPGVPGVPGV-GVPVGVPGVPGVPGVPGVPGVPGVPGVPGV 95
 QY 49 PGVGVTPAAAKA-----AKAAQGLVPGVAGVAGVAGVAGVAGVAGV 88
 Db 96 AGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 155
 QY 89 ---ASVGL---APGVGV---APGVGV---APGVGV---APGVGV---APGVGV 129
 Db 156 GVGVPVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 215
 QY 130 AKYGVTPAAAKA-----AKAAQGLVPGVAGVAGVAGVAGVAGV 172
 Db 216 AGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAGSGAG 272
 QY 173 PGVGLAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGV 200
 Db 273 PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV 296

RESULT 12
 PCT-US95-02772-4
 ; Sequence 4, Application US/08212237
 ; Patient No. 5606019
 ; GENERAL INFORMATION:
 ; APPLICANT: Cappello, Joseph
 ; TITLE OF INVENTION: Synthetic Proteins As Implantables
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
 ; STREET: Four Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/212,237
 ; FILING DATE: 11-MAR-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Rowland, Berttram I
 ; REGISTRATION NUMBER: 20,015
 ; REFERENCE/DOCKET NUMBER: A-58847/BIR
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-781-1989
 ; TELEFAX: 415-398-3249
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 832 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Protein
 US-08-212-237-4

Query Match 38.3%; Score 390; DB 1; Length 832;
 Best Local Similarity 57.1%; Pred. No. 5e-25;
 Matches 120; Conservative 10; Mismatches 34; Indels 46; Gaps 23;
 QY 3 GPGVGVGIPGV---AGVPGV---GVPGVG---GVPGVGSPERQAAAAAKA 50
 Db 57 GSGAGAGSPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 113
 QY 51 VCTPAAAAAKAAKAAQGLVPGVAGVAGVAGVAGVAGVAGVAGVAGVAGV 110
 Db 114 AGSGAGAGSGAGA---GSVPGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGV 152
 QY 111 GVAPALGPPEAAAAAKAAQGLVTPAAAAAKAAKAAQGLVPGVAGVAGVAGV 170
 Db 163 GV-PGVG---GAGAGSGA---GAGSGAGSGAGSGAGA---GSVPGVGV-PGVGV 207
 QY 171 VAPVGLAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGVAGV 200
 Db 208 V-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV 232

RESULT 13
 PCT-US95-02772-4
 ; Sequence 4, Application PC/TUS9502772
 ; GENERAL INFORMATION:
 ; APPLICANT: Protein Polymer Technologies, Inc.
 ; TITLE OF INVENTION: Synthetic Proteins As Implantables
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

```

; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/02772
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: FP-58847-1-PC/BIR
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; LENGTH: 832 amino acids
; STRANDS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-02772-4

```

```

Query Match 38.3%; Score 390; DB 5; Length 832;
Best Local Similarity 57.1%; Pred. No. 5e-25;
Matches 120; Conservative 10; Mismatches 34; Indels 46; Gaps 23;

Cy 3 GGVGVGGLPGVAGVGVG---GVPGVG---GVPGVGSIPFAQAQAAAKAAKYG 50
Db 57 GSGAGASGVPGVGVPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 113
Cy 51 VGTFAAAAKAAQAQFLVPGVGVAGVGVGVGVGVGVGVGVGVGVGVGVGVGV 110
Db 114 AGSGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 162
Cy 111 GVAPAGIPFAQAQAAAKAAKYGVTFAAAAKAAQAQFLVPGVGVAGVGVGVGV 170
Db 163 GV-PGVG---GAGAGSRA--GAGSAGAGAGAGAGAGAGAGAGAGAGAGAG 207
Cy 171 VAPGVGLAPGVAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 200
Db 208 V-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGV 232

```

```

RESULT 14
US-08-212-237-3
; Sequence 3, Application US/0821237
; Patent No. 5606019
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Synthetic Proteins As Implants
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/212.237
; FILING DATE: 11-Mar-1994
; CLASSIFICATION: A35
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58847/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 836 amino acids
; STRANDS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-212-237-3

```

```

Query Match 37.6%; Score 382; DB 1; Length 936;
Best Local Similarity 50.2%; Pred. No. 2.5e-24;
Matches 118; Conservative 8; Mismatches 41; Indels 68; Gaps 21;

Cy 3 GGVGVGGLPGVAGVGVG---GVPGVG---GVPGVGSIPFAQAQAAAKAAKYG 54
Db 45 GSGAGASGVPGV-GVPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 89
Cy 55 AAAAAAKAAQAQFLVPGVGVAGVGVGVGVGVGVGVGVGVGVGVGVGVGV 114
Db 90 GVGA-GAGSAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 141
Cy 115 AIGPFAQAQAAAKAAK---GVPGVG---GVPGVGSIPFAQAQAAAKAAK 145
Db 142 GVGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 200
Cy 146 AAAAAQFLVPGVGVAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 200
Db 201 GSGAGASGVPGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-PGVGV-P 248

```

```

RESULT 15
PCT-US95-02772-3
; Sequence 3, Application PC/TUS9502772
; GENERAL INFORMATION:
; APPLICANT: Protein Polymer Technologies, Inc.
; TITLE OF INVENTION: Synthetic Proteins As Implants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: PCT/US95/02772
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: FP-58847-1-PC/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

```

Search completed: April 24, 2001, 16:36:22
Job time: 302 sec

[illegible]

Q14233	PRELIMINARY;	PRT;	724 AA.
ID Q14233			
AC Q14233;	Q14238;		

```

Db 147 PGTVPGLPGAGFPGVGLPGVPGAGVFKPAGVPGAGFAGIPGVGPGGPGVPLG 206
Qy 181 PTKAPKLEGGYGLYTTKLPFGYGPVAGAGAGAGTGTGTGVPQAAAAAAXAKAF 240
Db 207 PTKAPKLEGGYGLYTTKLPFGYGPVAGAGAGAGTGTGTGVPQAAAAAAXAKAF 266
Qy 241 GAGAGVLPVGGAGVGPVGLPFGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
Db 267 GAGAGVLPVGGAGVGPVGLPFGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 326
Qy 301 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 360
Db 327 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 386
Qy 361 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 420
Db 387 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 446
Qy 421 AKAAYGVTPVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 480
Db 447 AKAAYGVTPVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 497
Qy 481 VGVAPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 540
Db 498 VGVAPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 547
Qy 541 GVPGLGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 600
Db 548 GVPGLGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 574
Qy 601 KYGAAGVGLGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 660
Db 575 KYGAAGVGLGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 634
Qy 661 GYGVGAGVGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 720
Db 635 GYGVGAGVGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 676

Query Watch      88 88; Score 3362; DB 4; Length 687;
      Local Similarity 90.38; Res No; 7e-154;
      Matches 660; Conservative 0; Mismatches 1; Indels 70; Gaps 3;
      1 GYGVGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 60
      27 GYGVGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 86
      61 PPGALVPGVADAAAAYKAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
      87 PPGALVPGVADAAAAYKAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 146
      121 SVTTCGVTPGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
      147 SVTTCGVTPGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 206
      181 PTKAPKLEGGYGLYTTKLPFGYGPVAGAGAGAGAGAGAGAGAGAGAGAGAG 240
      207 PTKAPKLEGGYGLYTTKLPFGYGPVAGAGAGAGAGAGAGAGAGAGAGAGAG 266
      241 GAGAGVLPVGGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 300
      267 GAGAGVLPVGGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 326
      301 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 360
      327 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 386
      361 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 420
      387 PGTGPGVPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 446
      421 AKAAYGVTPVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 480
      447 AKAAYGVTPVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 497
      481 VGVAPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 540
      498 VGVAPGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 547
      541 GVPGLGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 600
      548 GVPGLGVAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 574
      601 KYGAAGVGLGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 660
      575 KYGAAGVGLGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 634
      661 GYGVGAGVGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 720
      635 GYGVGAGVGLGAGVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPVGPV 676
```

```

RESULT 3
Q14235 PRELIMINARY; PRT; 687 AA.
Q14235;
DT 01-NOV-1996 (TRENDEL. 01, Created)
DT 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DT 01-OCT-2000 (TRENDEL. 15, Last annotation update)
DE ELASTIN.
GN ELN.
CC Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Primates; Catarrhini; Hominidae; Homo.
CC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=67274906; PubMed=3038460;
RA Indik Z., Yoon K., Morrow S.D., Cicella G., Rosenbloom J.,
RA Rosenbloom J., Ornstein-Goldstein N.;
RA "Structure of the 3' region of the human elastin gene: great abundance
ET of Alu repetitive sequences and few coding sequences.";
RL Connect. Tissue Res. 16:197-211 (1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=28366; PubMed=3038501;
RA Indik Z., Ornstein-Goldstein N., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Palonen L., Rosenbloom J.,
RA "Alternative splicing of human elastin mRNA indicated by sequence
ET analysis of cloned genomic and complementary DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684 (1987).
```

QY	721	CUGKACGRKK	731
DB	677	CUAGACRKK	687
 RESULT 4			
ID	OL5336	PRELIMINARY;	PRT; 635 AA.
AC	OL5336	PRELIMINARY;	PRT; 635 AA.
DT	01-JAN-1998	(TREMELZel. 05, Created)	
DT	01-JAN-1998	(TREMELZel. 05, Last sequence update)	
DE	01-OCT-2000	(TREMELZel. 15, Last annotation update)	
DE	ELASIN (FRAGMENT)		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCB1_Taxid=9606;		
RN	SEQUENCE FROM N.A.		
RP	LI D.Y., Toland A.E., Boak B.B., Atkinson D.L., Ensing G.J.,		
RA	Morris C.A., Keating M.T.;		
RE	Hum. Mol. Genet. 6:0-0(1997).		
DR	EMBL; U93037; AAB55221.1; JOINED		
DR	EMBL; U93034; AAB55221.1; JOINED		
DR	EMBL; U93035; AAB55221.1; JOINED		
DR	EMBL; U93036; AAB55221.1; JOINED		
DR	EMBL; U93037; AAB55221.1; JOINED		
DR	EMBL; U93038; AAB55221.1; JOINED		
DR	EMBL; U93039; AAB55221.1; JOINED		
DR	EMBL; U93040; AAB55221.1; JOINED		
DR	EMBL; U93041; AAB55221.1; JOINED		
DR	EMBL; U93042; AAB55221.1; JOINED		
DR	EMBL; U93043; AAB55221.1; JOINED		
DR	EMBL; U93044; AAB55221.1; JOINED		
DR	EMBL; U93045; AAB55221.1; JOINED		
DR	EMBL; U93046; AAB55221.1; JOINED		
DR	EMBL; U93047; AAB55221.1; JOINED		
DR	EMBL; U93048; AAB55221.1; JOINED		
DR	EMBL; U93049; AAB55221.1; JOINED		
DR	EMBL; U93050; AAB55221.1; JOINED		
DR	EMBL; U93051; AAB55221.1; JOINED		
DR	EMBL; U93052; AAB55221.1; JOINED		
DR	EMBL; U93053; AAB55221.1; JOINED		
DR	EMBL; U93054; AAB55221.1; JOINED		
DR	EMBL; U93055; AAB55221.1; JOINED		
DR	EMBL; U93056; AAB55221.1; JOINED		
DR	EMBL; U93057; AAB55221.1; JOINED		
DR	EMBL; U93058; AAB55221.1; JOINED		
DR	EMBL; U93059; AAB55221.1; JOINED		
DR	EMBL; U93060; AAB55221.1; JOINED		
DR	EMBL; U93061; AAB55221.1; JOINED		
DR	EMBL; U93062; AAB55221.1; JOINED		
DR	EMBL; U93063; AAB55221.1; JOINED		
DR	EMBL; U93064; AAB55221.1; JOINED		
DR	EMBL; U93065; AAB55221.1; JOINED		
DR	EMBL; U93066; AAB55221.1; JOINED		
DR	EMBL; U93067; AAB55221.1; JOINED		
DR	EMBL; U93068; AAB55221.1; JOINED		
DR	EMBL; U93069; AAB55221.1; JOINED		
DR	EMBL; U93070; AAB55221.1; JOINED		
DR	EMBL; U93071; AAB55221.1; JOINED		
DR	EMBL; U93072; AAB55221.1; JOINED		
DR	EMBL; U93073; AAB55221.1; JOINED		
DR	EMBL; U93074; AAB55221.1; JOINED		
DR	EMBL; U93075; AAB55221.1; JOINED		
DR	EMBL; U93076; AAB55221.1; JOINED		
DR	EMBL; U93077; AAB55221.1; JOINED		
DR	EMBL; U93078; AAB55221.1; JOINED		
DR	EMBL; U93079; AAB55221.1; JOINED		
DR	EMBL; U93080; AAB55221.1; JOINED		
DR	EMBL; U93081; AAB55221.1; JOINED		
DR	EMBL; U93082; AAB55221.1; JOINED		
DR	EMBL; U93083; AAB55221.1; JOINED		
DR	EMBL; U93084; AAB55221.1; JOINED		
DR	EMBL; U93085; AAB55221.1; JOINED		
DR	EMBL; U93086; AAB55221.1; JOINED		
DR	EMBL; U93087; AAB55221.1; JOINED		
DR	EMBL; U93088; AAB55221.1; JOINED		
DR	EMBL; U93089; AAB55221.1; JOINED		
DR	EMBL; U93090; AAB55221.1; JOINED		
DR	EMBL; U93091; AAB55221.1; JOINED		
DR	EMBL; U93092; AAB55221.1; JOINED		
DR	EMBL; U93093; AAB55221.1; JOINED		
DR	EMBL; U93094; AAB55221.1; JOINED		
DR	EMBL; U93095; AAB55221.1; JOINED		
DR	EMBL; U93096; AAB55221.1; JOINED		
DR	EMBL; U93097; AAB55221.1; JOINED		
DR	EMBL; U93098; AAB55221.1; JOINED		
DR	EMBL; U93099; AAB55221.1; JOINED		
DR	EMBL; U93100; AAB55221.1; JOINED		
DR	EMBL; U93101; AAB55221.1; JOINED		
DR	EMBL; U93102; AAB55221.1; JOINED		
DR	EMBL; U93103; AAB55221.1; JOINED		
DR	EMBL; U93104; AAB55221.1; JOINED		
DR	EMBL; U93105; AAB55221.1; JOINED		
DR	EMBL; U93106; AAB55221.1; JOINED		
DR	EMBL; U93107; AAB55221.1; JOINED		

[illegible][illegible]


```

QY 204 --YGPQGVAGA-----AGTASY-PGTGVTGFPQAAAAAANKAAKAFKRGAGAGVDPV 251
DB 1044 GYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 1103
QY 252 GGNAGVPGVPGVPLGIGLAGVGTAAATAAANKAAKAYG-AAAGLVFG--GPGFPGVYV 308
DB 1104 PFGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYV 1163
QY 309 GYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 367
DB 1164 GYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 1217
QY 368 IPTYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 420
DB 1218 LGNGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 1277
QY 421 AKAAKYGVTAAATAAANKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 480
DB 1278 LDITTDGAGGTITISEELTISGAGS-----GSGPG-GAGPG-GYVGP-GSGPG-GYVGP 1328
QY 481 VGVAGVGPVAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 540
DB 1329 -GSGPG-----GVFPGAGGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 1366
QY 541 GYVPLGVAGVPGFAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 595
DB 1367 GSGGPGGAGPG-YVGP-GYVGP-GYVGP-GYVGP-GYVGP-GYVGP-GYVGP-GYVGP-GYVGP 1424
QY 596 ---AKAAKYGVTAAATAAANKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK 652
DB 1425 YGPGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 1480
QY 653 GLGGLGVGLGVGPVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 708
DB 1481 GPGGPGGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 1538
QY 709 GYVPLGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 720
DB 1539 GYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGP 1550

RESULT 14
Q9NHM4
ID 05NHM4 PRELIMINARY: PRT: 2249 AA.
AC 05NHM4
DT 01-OCT-2000 (TREMELrel. 15, Created)
DT 01-OCT-2000 (TREMELrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE FLAGELLIFORM SILK PROTEIN (FRAGMENT).
GN FLAG.
OS Nephila clavipes (Orb spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OX Araneomorphae; Eutelesynae; Araneidae; Tetragnathidae; Nephila.
NXI_TaxID=6915;
RN [1]
RP SEQUENCE FROM N.A.
RA Hayashi C.Y., Lewis R.V.;
RT "Molecular architecture and evolution of a modular spider silk protein
gene."; 287-1477-1479 (2000).
DB EMBL:AB218621; ACP36090.1;
FT NCON_TER 2249 2249
SQ SEQUENCE 2249 AA; 174871 MW; 88CB168A147CDA0 CRC64;

Query Match 19.1%; Score 723; DB 5; Length 2249;
Best Local Similarity 35.6%; Pred. No. 2.6e-27;
Matches 268; Conservative 22; Mismatches 325; Indels 138; Gaps 43;
QY 2 GYVPLGVAGVPGFAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
DB 1541 GPGGSGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAGPGYVGPQGGAG 1595

```

SQ SEQUENCE 907 AA; 72392 MW; A0864ECT08740A00 CRC64;

Query March 18.4%; Score 698; DP 5; Length 907;
 Best Local Similarity 33.2%; Pred. No. 1.8e-26;
 Matches 261; Conservative 29; Mismatches 309; Indels 186; Gaps 42;

QY 2 GVEGAIQGVG-PGVSFVFGAGLALGAGGAGPQK-----PLKPVPGGILAGAGLGG 54
 Db 131 GPGSGSGGPGGPGSGPGAGPGGVGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGG 189
 QY 55 AFPAVTPGALVPGGVADNAAAYKAKAGAGLGGVPGVGGVGGVGSAGAVPQ----- 105
 Db 190 AGPGGAGPGGAGPGGPGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 249
 QY 106 -----PQAGPKPKVPGVGLGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 153
 QY 250 DGPITISBELIISAG-----GSGPG-----GAGPGGVPGGSGPG-VGPG----- 293
 QY 154 PGVGAGAFGPGVPGFPGPGQVPLAYIPKAPLPGGCLPYTKLPGVGGPGVAGAA 213
 Db 294 PG-----GVPG-GSGPGF-GVG-----PGGAGSGPYGPG-----GSGPGGAGGAG 330
 QY 214 GKAG-YPTGTGVPQAAAAAKAAKAGAGAGLVPGCVGAGVPGVPGAPIGSIGIAGY 272
 Db 331 GPGGAYGPGSGTGP-GSGSGPGGAGGPGTGP-----EPCGAGGPGVPGGAGPGTGP 385
 QY 273 TPGAAAAAANAANAANKYGAAGAGLVPGSG--FGPGVGVWP-GRGVP-----GVGVP 325
 Db 385 GGP-----YGGSGSGPGVPGSGSYGPGSGSYGPGSGSYGPGVPGVPGGPGSGAG 431
 QY 326 PVPVPGAGIPGAAPVGVGVSFAAKAAKAAKAGPAGVGVGGVPGVPGVAGGPGPG 385
 Db 432 PPGGPGVPGSGSGPGGPGGSGPGGPG-----YGP-GRAGPGGVPGSGSGPGVPG 485
 QY 386 GGI-PTVAGVPSVG-GVPGVGVPGVGVSPERAAAAAKAAKYGVPVPGAAAAAKAA 443
 Db 486 GGTGPGAGSGPGGTGPGGSGPGGAGGAGGAGGPGTGPGGSGPGGTGPG----- 531
 QY 444 DFLGVPGVAGVGVGVPVGVAGVAGLVAGVPGVAGVGVGVPAGVPGVPGVAAA 503
 Db 532 -GSGPG-GTGP-GTGP-GTGP-GTGP-GTGP-GTGP-GTGP-GTGP-GTGP-GTGP 583
 QY 504 SNAKYAAKQLRRAALGAGLPGVGVGPGVGLGVAGKVPGLGVAGVPGV----- 553
 Db 584 GPGGVGPGSGPGGTGPGSGSPG-----GYGPGSGAGGTGPGGAGGAGGAGSGGSGG 640
 QY 554 ----TGAGADGVARSLEPHE-----GPFSSQHLPTSPVPPVPGNLAARAA 599
 Db 641 AGSGGAGSGSGVSGSGGTITFDLOITIDAGDPTISEELTISGAGSGPGGAGPG 700
 QY 600 AYGARVPGVGGILGALGPGVPGVGVGPGPAAAAAKAAAKAAAGFLVAGLGGV 659
 Db 701 GPGSGPGVPGVPGVPGVPGVPG-VG-----GSGPGVGS-----GSGPGVPG 742
 QY 660 GGLGVPGVGGTGTTPAAAAKAAKYGAGLGVGLGGAG-----PLGQVAGRGGLSP 715
 Db 743 GGVG-PGSGSGSGVGPSPG-----YGPGSGSGGTGPGGSGGTGPGSGGTG 791
 QY 716 PPGGA 720
 Db 792 -PGCA 795

Search completed: April 24, 2001, 16:40:24
Job time: 414 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: April 24, 2001, 16:38:14 ; Search time 115.25 Seconds
(without alignments)
99.695 Million cell updates/sec

Title: US-09-340-736-2
Perfect score: 1017
Sequence: 1 PFGEVGVGVPVGVGVGV.....GVAPGVGVGVGVAPGIP 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04

Listing first 45 summaries

Database : A Genesep 0401.*
1: /SIDSI/gcgdata/genesep/AA1980.DAT.*
2: /SIDSI/gcgdata/genesep/AA1981.DAT.*
3: /SIDSI/gcgdata/genesep/AA1982.DAT.*
4: /SIDSI/gcgdata/genesep/AA1983.DAT.*
5: /SIDSI/gcgdata/genesep/AA1984.DAT.*
6: /SIDSI/gcgdata/genesep/AA1985.DAT.*
7: /SIDSI/gcgdata/genesep/AA1986.DAT.*
8: /SIDSI/gcgdata/genesep/AA1987.DAT.*
9: /SIDSI/gcgdata/genesep/AA1988.DAT.*
10: /SIDSI/gcgdata/genesep/AA1989.DAT.*
11: /SIDSI/gcgdata/genesep/AA1990.DAT.*
12: /SIDSI/gcgdata/genesep/AA1991.DAT.*
13: /SIDSI/gcgdata/genesep/AA1992.DAT.*
14: /SIDSI/gcgdata/genesep/AA1993.DAT.*
15: /SIDSI/gcgdata/genesep/AA1994.DAT.*
16: /SIDSI/gcgdata/genesep/AA1995.DAT.*
17: /SIDSI/gcgdata/genesep/AA1996.DAT.*
18: /SIDSI/gcgdata/genesep/AA1997.DAT.*
19: /SIDSI/gcgdata/genesep/AA1998.DAT.*
20: /SIDSI/gcgdata/genesep/AA1999.DAT.*
21: /SIDSI/gcgdata/genesep/AA2000.DAT.*
22: /SIDSI/gcgdata/genesep/AA2001.DAT.*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1017	100.0	201	19 W46316	Non-natural polype
2	709.5	69.8	598	20 Y01302	Human tropoelastin
3	709.5	69.8	598	21 Y69069	Amino acid sequenc
4	709.5	69.7	560	20 Y01303	Human tropoelastin
5	707.5	69.6	712	21 B08630	Amino acid sequenc
6	707.5	69.6	730	21 B08631	Fusion protein com
7	700	68.8	730	21 W46315	Human elastin cont
8	700	68.8	730	21 W46315	Human elastin cont
9	700	68.8	733	15 P56553	Synthetic human el
10	700	68.8	733	20 Y01301	Amino acid sequenc
11	683.5	67.2	571	21 Y69071	Amino acid sequenc

12	670	65.9	515	21 Y69135	Amino acid sequenc
13	531.5	52.3	745	20 Y31682	Alanine-containing
14	455.5	44.8	148	20 Y31681	Alanine-containing
15	446.5	43.9	294	9 P82484	Tropoelastin, Gal
16	401	39.4	988	16 R80253	Polymer:SELP7, Sy
17	397.5	39.1	972	16 R80255	Polymer:SELP4, Sy
18	392	38.5	1056	16 R80254	Polymer:SELP3, Sy
19	393.5	38.5	1824	16 R80256	Polymer:SELP8, Sy
20	382	37.8	936	16 R80257	Polymer:SELP9, Sy
21	382	37.6	936	16 R80261	Polymer:SELP0, Sy
22	379.5	37.3	1412	19 W53519	Amino acid sequenc
23	379.5	37.3	1413	9 P82957	ESBI protein compr
24	379.5	37.3	1413	14 R41008	ESBI multimeric pr
25	379.5	37.3	1413	18 W26343	ESBI multimeric el
26	379.5	37.3	1464	21 Y78278	ESBI protein sequ
27	376.5	37.0	281	19 W47836	Protein polymer ad
28	376.5	37.0	287	13 W47837	Protein polymer ad
29	376.5	37.0	287	13 W47837	Protein polymer ad
30	376.5	37.0	287	13 W47837	Protein polymer ad
31	371.5	36.5	2257	9 P82661	SELP3 protein comp
32	371.5	36.5	2257	14 R41012	SELP3 multimeric p
33	371.5	36.5	2257	18 W26347	SELP3 synthetic el
34	371.5	36.5	2257	19 W53523	Amino acid sequenc
35	371.5	36.5	2257	21 Y78282	SELP3 amino acid s
36	368.5	36.2	378	17 W09219	SELP0K polymer, S
37	368.5	36.2	378	19 W53545	Amino acid sequenc
38	368.5	36.2	378	21 Y51889	Plasmid:PT0375 pr
39	368.5	36.2	1002	17 W09218	SELP0K polymer, S
40	368.5	36.2	1002	17 W09218	Amino acid sequenc
41	368.5	36.2	1002	17 W09218	Plasmid:PT0364nc
42	365.5	35.0	2055	9 P82860	SELP2 protein comp
43	365.5	35.0	2055	14 R41011	SELP2 multimeric p
44	366.5	36.0	2055	18 W26346	SELP2 synthetic el
45	366.5	36.0	2055	19 W53522	Amino acid sequenc

ALIGNMENTS

RESULT 1
W46316
ID W46316 standard; protein: 201 AA.
XX
AC W46316;
XX
DT 23-JUL-1998 (first entry)
XX
DE Non-natural polypeptide MFU-2 sequence.
XX MFU-1, minimal functional unit; elastin; human, fibrous protein;
KW beta-sheet, coating, wound dressing, MFU
XX Homo sapiens.
XX
XX
XX W09805695-R2.
XX
XX 12-FEB-1998.
XX
XX 07-AUG-1997; 97W0-CA00560.
XX
PR 07-AUG-1997; 97W0-0911364.
XX
PR 07-AUG-1996; 96W0-0023552.
XX
PA (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX (PROT-) PROTEIN SPECIALTIES LTD.
XX
PI Keeley FW, Rothstein A, Rothstein SJ;
XX
XX WPI, 1998-145551/13.
XX
PT New non-natural polypeptide with multiple beta-sheet, beta-tur
PT structures - particularly based on human elastin, useful for coating

[illegible]

Db 461 vaaakaakvaakadlraaaglgagipglvgvgvlgvgaavglvgvgvgaav 520
 QY 125 -----AAKXAXY-----GVGP-----AAAKAAKAAAFGLVVG 156
 Db 521 pgaataakayavgyvglgaglgvgvgvgvgvgvgvgvgvgvgvgvgvgvg 160
 QY 157 vgvagvgvgv-----AGVGAGVG-----GLAGVGAGVG-----GVGAGVG 193
 Db 581 aglg-glgvg 639
 QY 194 GVAP 197
 Db 640 gsp 643

RESULT 5

B08630 ID B08630 standard; peptide; 712 AA.

AC B08630;

20-DEC-2000 (first entry)

Amino acid sequence of a human elastin polypeptide.

XX Tropoelastin; elastin; elastic fibre; smooth muscle cell proliferation;
 KW smooth muscle cell differentiation; smooth muscle cell migration;
 KW smooth muscle cell function; atherosclerosis; restenosis; aneurysm;
 KW vascular bypass graft stenosis; transplant arteriopathy; dissection;
 KW SVAS; hypertension; transplant arteriopathy.

XX Homo sapiens.

XX WO200005068-A2.

XX 31-AUG-2000.

XX 28-FEB-2000; 2000WO-US02526.

XX 26-FEB-1999; 99US-025217.

XX (UTAH) UNIV UTAH RES FOUND.

XX Keating MT, Li DY;

XX WPI; 2000-533134/48.

XX Elastin based compositions useful for treating atherosclerosis,
 PT restenosis, vascular bypass graft stenosis, transplant arteriopathy,
 PT aneurysm, dissection SVAS and/or hypertension -

Example 3; Page 46; 79pp; English.

XX The present sequence represents a human elastin. Peptides derived
 CC from elastin are used in compositions of the invention. The protein is used
 CC in compositions that are potent regulators of smooth muscle cell
 CC migration in vivo. The elastin-based compositions comprise at least one
 CC elastic fibre, elastins, tropoelastins (or fragments of them) which have
 CC biological activities comprising: inhibiting the proliferation of smooth
 CC muscle cells in vivo; stimulating the differentiation of smooth muscle
 CC cell in vivo; and regulating the migration of smooth muscle cells in
 CC vivo. The compositions may be used for the prophylaxis or treatment
 CC of a disorder characterized by diminished capacity to regulate smooth
 CC muscle cell function such as atherosclerosis, restenosis, vascular
 CC bypass graft stenosis, transplant arteriopathy, aneurysm and/or
 CC dissection. Disorders which may be treated also include SVAS (undefined),
 CC hypertension, and transplant arteriopathy.

XX Sequence 712 AA;

Query Match 69.6%; Score 707.5; DB 21; Length 712;

Best Local Similarity 55.7%; Pred. No. 8.8e-45;
 Matches 166; Conservative 6; Mismatches 19; Indels 107; Gaps 8;

QY 1 FPGVGVGIFGVAGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVG 60

Db 404 tpgvg 463

QY 61 AAARAAQF-----GLVFGVGAGVGAGVGAGVGAGVGAGVGAGVGAGVG 114

Db 464 aaakaagfallnlaglvgvgvgvgvgvgvgvgvgvgvgvgvgvgvgvgvg 523

QY 115 AIGPPAQARA----- 124

Db 524 gtpgvgvaakaakaaakaaaglgagipglvgvgvgvgvgvgvgvgvgvgvg 583

QY 125 -----AAAKAKGVGP----- 149

Db 534 pfgvgvgalaaakaaakaaagvgvgvgvgvgvgvgvgvgvgvgvgvgvgvg 643

QY 150 OFGLVFGVGAGVGAGVGAGVGAGVGAGVGAGVGAGVGAGVGAGVGAGVG 197

Db 644 qfglvgaglg-glgvg-glgvg-pgvg---glgglppaaakaakaygvaarpfglsp 695

RESULT 6

B08631 ID B08631 standard; peptide; 730 AA.

XX B08631;

20-DEC-2000 (first entry)

Fusion protein comprising human elastin and c-myc.

XX Tropoelastin; elastin; elastic fibre; smooth muscle cell proliferation;
 KW smooth muscle cell differentiation; smooth muscle cell migration;
 KW smooth muscle cell function; atherosclerosis; restenosis; aneurysm;
 KW vascular bypass graft stenosis; transplant arteriopathy; dissection;
 KW SVAS; hypertension; transplant arteriopathy.

XX Synthetic.

XX Homo sapiens.

XX Unidentified.

XX WO200005068-A2.

XX 31-AUG-2000.

XX 28-FEB-2000; 2000WO-US02526.

XX 26-FEB-1999; 99US-025217.

XX (UTAH) UNIV UTAH RES FOUND.

XX Keating MT, Li DY;

XX WPI; 2000-533134/48.

XX Elastin based compositions useful for treating atherosclerosis,
 PT restenosis, vascular bypass graft stenosis, transplant arteriopathy,
 PT aneurysm, dissection SVAS and/or hypertension -

Example 3; Page 48; 79pp; English.

XX The present sequence represents a fusion protein, comprising human
 CC elastin and c-myc, preceded by a His tag. The protein is used in
 CC compositions of the invention. The specification describes elastin
 CC based compositions that are potent regulators of smooth muscle cell
 CC migration in vivo. The elastin-based compositions comprise at least one
 CC elastic fibre, elastins, tropoelastins (or fragments of them) which have
 CC biological activities comprising: inhibiting the proliferation of smooth
 CC muscle cells in vivo; stimulating the differentiation of smooth muscle
 CC cell in vivo; and regulating the migration of smooth muscle cells in
 CC vivo. The compositions may be used for the prophylaxis or treatment
 CC of a disorder characterized by diminished capacity to regulate smooth
 CC muscle cell function such as atherosclerosis, restenosis, vascular
 CC bypass graft stenosis, transplant arteriopathy, aneurysm and/or
 CC dissection. Disorders which may be treated also include SVAS (undefined),
 CC hypertension, and transplant arteriopathy.

[illegible][illegible]

SQ Sequence 294 AA:

```
Query Match      43.9%; Score 446.5; DS 9; Length 294;
Best Locality    54.1%; Freq. Ref. 3.2e-26;
Matches 120; Conservative 11; Mismatches 44; Indels 47; Gaps 16;

| QY 2 PGFGVGGIPGVAGVGVG---GVFGVG---GVFGVG-----SPDAQAAAKA 45
|   || || || || || || || || || || || || || || || || || || ||
| Db 4 PGVGPVGVGVGV-GVPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 62
|   || || || || || || || || || || || || || || || || || || ||
| QY 47 AKYGVGTPAAAKAAAKAAAGSLVPGV-GVAPGV-GVAPGV-GVAPGV-GVAPGV 101
|   || || || || || || || || || || || || || || || || || || ||
| Db 63 AKYgag-----GLAPGVGLAPGVGLAPGVGLAPGVGLAPGVGLAPGVGLA 106
|   || || || || || || || || || || || || || || || || || || ||
| QY 102 PGV-GVAPGVGVAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 159
|   || || || || || || || || || || || || || || || || || || ||
| Db 107 PGVGLAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 162
|   || || || || || || || || || || || || || || || || || || ||
| QY 160 APGVGVAPGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGVGV 201
|   || || || || || || || || || || || || || || || || || || ||
| Db 163 tpgvgvvp--slvpgvgv--pgvgvlpvggi--pgvgvqgirkp 200
|   || || || || || || || || || || || || || || || || || || ||
```

Search completed: April 24, 2001, 16:38:15
Job time: 415 sec

Result	No.	Score	Query Match	Length	DB	ID	Description
1	562.5	65.1	792	1	ERABU		elastin precursor,
2	545.5	53.6	770	2	S59623		tropoelastin - she
3	485.5	47.7	747	1	ERABO		elastin precursor,
4	474	46.6	784	2	A26601		elastin precursor
5	412	40.5	864	1	ERART		elastin precursor
6	379	37.3	860	1	ENMS		elastin precursor
7	346.5	34.1	907	2	A35560		sporozoite surface
8	243.5	24.5	136	2	A25883		elastin - bovine (
9	237.5	23.7	137	2	A25885		elastin - bovine (
10	230	22.6	148	2	P70807		hypothetical glyci
11	223	21.9	1901	2	P70806		hypothetical glyci
12	216.5	21.3	416	1	SKXLAG		dermal gland prote
13	214.5	21.1	445	2	A27263		50K spicule matrix
14	209.5	20.6	641	1	Q06351		nuclear antigen EB
15	205.5	20.2	1660	2	A70869		hypothetical glyci
16	198.5	19.5	713	1	UNMS		period clock prote
17	195.5	19.2	419	2	G70602		hypothetical prote
18	192.5	18.9	627	2	A44112		spidroin 2, dragli
19	192	18.9	2639	2	T31328		fibroin - Chinese
20	188	18.5	743	1	S33779		collagen alpha 1(V
21	187.5	18.4	1306	2	A70934		hypothetical glyci
22	186.5	18.3	744	2	S52835		collagen alpha 1(V
23	186.5	18.2	746	2	S52835		collagen alpha 1(V
24	185.5	18.2	746	2	T32455		collagen alpha 1(V
25	185	18.2	744	1	A34246		collagen alpha 1(V
26	185	18.2	744	1	S33998		collagen alpha 2(I
27	184.5	18.1	1763	2	S16356		collagen alpha 2(I
28	184	18.1	498	2	G70720		hypothetical glyci
29	183.5	18.0	413	2	E70561		proable PE protei

[illegible]

Matches	62;	Conservative	2;	Mismatches	9;	Indels	29;	Gaps	4;
QY	52	GTPAAA--AAKAAAKAQAQGLVPGVGVA	PGVGVA	PGVGVA	PGVGVA	PGVGLAPGVGVA	PGVGVAPG	VAPG	109

DB 479 AKAARKGAGGAGAGLGLPGAVPGALPGAVPAVPGAGVPGAGTAAAAAATAAK 538

QY 175 VGLAPGVGVAP-----GVGVAPG-----YGVAPAIGP 201

	Matches	62; Conservative	2; Mismatches	9; Indels	29; Gaps	4;
QY	52	GTPAAA--AAKAAAKAQAQGLVPGVGVAFCGVGVAPGCVGLAPGCVGVAPG	109			

[illegible][illegible]

Query Match 21.1%; Score 214.5; DB 2; Length 445;
Best Local Similarity 34.5%; Pred. No. 1.3e-06;
Matches 77; Conservative 13; Mismatches 54; Indels 79; Gaps 14;

[illegible]

RESULT 15

C;Superfamily: collagen alpha 1(IV) chain

Arch completed: April 24, 2001, 16:41:54
Time: 459 sec

Result	Score	Query Match	Length	DB	TD	Description
1	696.5	68.5	730	1	ELS_HUMAN	P15502 homo sapien
2	485.5	47.7	740	1	ELS_BOVIN	P04985 bos taurus
3	477	46.9	750	1	ELS_CHICK	P04935 gallus gall
4	412	40.5	864	1	ELS_RAT	Q9372 zatus norv
5	379	37.3	860	1	ELS_MOUSE	P54320 mus musculus
6	256.5	25.2	5363	1	PEOR_BOVMO	P05790 boobyx mori
7	223	21.9	1301	1	Y208_MYCTU	Q53553 mycobacteri
8	216.5	21.3	439	1	X22_XENLA	P17437 xenopus lae
9	214.5	21.1	445	1	SP50_STIRPU	P11994 strongyloce
10	209.5	20.6	641	1	ENLI_EBV	P03211 Epstein-Bar
11	180.6	16.6	1471	1	ENLI_MOUSE	P03211 Epstein-Bar
12	198.5	19.2	1677	1	ENLI_MOUSE	P03211 Epstein-Bar
13	193.5	18.9	607	1	ENLI_MOUSE	P03211 Epstein-Bar
14	192.5	18.9	627	1	SP02_MOUSE	P03211 Epstein-Bar
15	188	18.5	743	1	CA18_MOUSE	P46604 nephila gla
16	186.5	18.3	743	1	CA18_MOUSE	P00780 homo musculus
17	185	18.2	744	1	CA18_EBKIT	P27658 homo sapien
18	184.5	18.1	1763	1	CA24_ASCARU	P14262 oryctolagu
19	184	18.1	498	1	Y118_MYCTU	P27393 ascaris suu
20	182.5	17.9	278	1	Y186_CAEEL	Q50615 mycobacteri
21	180.5	17.7	157	1	C063_LOCOM	P34391 caenorhabdi
22	180.5	17.7	1659	1	CA14_HUMAN	P45584 locusta mig
23	180	17.7	1373	1	CA21_MOUSE	P02462 homo sapien
24	179.5	17.6	384	1	GA21_PEPHY	Q01149 mus musculus
25	177	17.4	1883	1	Y23D_HUMAN	P09789 petunia byb
26	177	17.4	1885	1	CA5D_HUMAN	P00268 homo sapien
27	176.5	17.4	308	1	C040_CAEEL	P29400 homo sapien
28	176	17.3	148	1	Y065_MYCTU	P34804 caenorhabdi
29	175	17.2	148	1	Y065_MYCTU	P34804 caenorhabdi
30	173	17.2	148	1	Y065_MYCTU	P34804 caenorhabdi
31	173	17.2	148	1	Y065_MYCTU	P34804 caenorhabdi
32	174	17.1	1464	1	CA21_CAEEL	P45584 locusta mig
33	174	17.1	1603	1	CA11_HUMAN	P02452 homo sapien
34	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
35	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
36	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
37	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
38	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
39	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
40	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
41	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
42	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
43	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
44	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
45	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
46	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
47	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
48	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
49	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
50	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
51	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
52	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
53	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
54	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
55	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
56	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
57	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
58	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
59	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
60	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
61	173	17.0	152	1	CA11_HUMAN	P00792 homo sapien
62						

[illegible]

Query Match	21.9%	Score 223;	DB 1;	Length 1901;
Best Local Similarity	32.9%;	pred. No. 7.9e-06;		
Matches 77:	Conservative	7;	Mismatches 114;	Indels 36;
	Gaps	4;		

[illegible]

[illegible]

Query Match 19.2%; Score 195.5; DB 1; Length 407;
Best Local Similarity 35.6%; Pred. No. 8.2e-05;

GenCore version 4.5
Copyright (c) 1995 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:40:24 ; Search time 125.5 seconds
(without alignments)
187.719 Million cell updates/sec

Title: US-09-340-736-2

Percent score: 107

Sequence: 1 FFGFGVGVGIRGAGVGV.....GVAPGVGAGVGVAPLGP 201

Scoring table:

BLOSUM62

Gapop 10.0 ; Gapext 0.5

374700 segs. 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREML15.*

1: SP_Archaea.*

2: SP_Bacteria.*

3: SP_Eukarya.*

4: SP_Humans.*

5: SP_Invertebrates.*

6: SP_Mammals.*

7: SP_MHC.*

8: SP_Organelle.*

9: SP_Phage.*

10: SP_Plant.*

11: SP_Rodent.*

12: SP_Unclassified.*

13: SP_Vertebrate.*

14: SP_Virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
1	709.5	59.8	774	4	Q14233	Q14233	homo sapien
2	700	68.8	777	4	Q14234	Q14234	homo sapien
3	655.5	64.5	602	4	Q15337	Q15337	homo sapien
4	654.5	64.4	635	4	Q15336	Q15336	homo sapien
5	606	59.6	657	4	Q14235	Q14235	homo sapien
6	498	49.0	650	6	Q28099	Q28099	bos taurus
7	490	48.2	679	6	Q28097	Q28097	bos taurus
8	490	48.2	707	6	Q28098	Q28098	bos taurus
9	477	47.9	666	6	Q28096	Q28096	bos taurus
10	371.5	36.5	258	4	Q98K93	Q98K93	homo sapien
11	341.5	33.5	258	5	Q98K92	Q98K92	homo sapien
12	291.5	28.7	1729	5	Q98K17	Q98K17	homo sapien
13	287.5	28.7	1039	5	Q98K96	Q98K96	homo sapien
14	260	25.6	988	5	Q17434	Q17434	neophila cla
15	249.5	24.5	76	6	Q28100	Q28100	bos taurus
16	240.5	23.6	120	14	Q88914	Q88914	thermoprote
17	237	23.3	1334	2	Q98K95	Q98K95	streptomyce
18	235.5	23.2	117	14	Q88913	Q88913	thermoprote
19	230	22.6	1489	2	Q53559	Q53559	mycobacteri

ALIGNMENTS

RESULT 1

Q14233 PRELIMINARY; PRG; 724 AA.
AC Q14233; Q14238; (TRENDEL. 01, Created)
DT 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DT 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DT 01-OCT-2000 (TRENDEL. 15, Last annotation update)
DE ELASTIN.
GN ELN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NC Taxid=9606;
RN 1
RX MEDLINE=6724966; PubMed=3039501;
RA Rosenbloom J., Ornstein-Goldstein N.;
RT "Alternative splicing of human elastin mRNA indicated by sequence
Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987)."
RN [3]
RX MEDLINE=8156138; PubMed=2831431;
RA Razio M.J., Olsen P., Kuvshinov H., Chu M.L., Davidson J.M.,
RT "Isolation and characterization of human elastin cDNAs, and age-
associated variation in elastin gene expression in cultured skin
fibroblasts".
RN Lab. Invest. 58:270-277(1988).
DR EMBL; M17282; AAC98394.1;
DR EMBL; M16983; AAC98394.1; JOINED.
DR EMBL; M17265; AAC98394.1; JOINED.

Q15985 araneus dia
Q01927 phytophthor
Q01914 phytophthor
Q01914 phytophthor
Q53215 mycobacteri
Q02402 pinctada fi
Q01916 phytophthor
Q05499 chlamydomon
Q05589 mycobacteri
Q23018 arabidopsis
Q76786 antheraea p
Q15988 araneus dia
Q15988 streptomyce
Q21436 neophila cla
Q04844 neophila cla
Q97548 drosophila
Q46171 neophila cla
Q53775 mycobacteri
Q17163 brugia mala
Q53435 mycobacteri
Q97210 macaca muia
Q9K15 streptomyce
Q17374 mycobacteri
Q94236 mycobacteri
Q43688 vigna ungui

2	PRELIMINARY; PRT; 757 BA.
1	AC Q14234
0	DT 01-NOV-1996 (TReMELrel. 01, Created)
0	DT 01-NOV-1996 (TReMELrel. 01, Last sequence update)
0	DT 01-OCT-2000 (TReMELrel. 15, Last annotation update)
0	DE ELASTIN.
0	EN ELN.
0	OS Homo sapiens (Human).
0	CS Bakaryoua; Baccharis; Chordata; Craniata; Vertebrata; Euteleostomi;
0	OC Eukaryota; Euteleostomi; Euteleostomi; Euteleostomi; Euteleostomi;
0	CC NCBI_TaxID=9606;
0	EN [1]
0	RP SEQUENCE FROM N.A.
0	RX MEDLINE=87274906; PubMed=3038460;
0	RA Indik Z., Yoon K., Morrow S.D., Cicilia G., Rosenbloom J.,
0	RN Rosenbloom J., Orstein-Goldstein N.;
0	RT A structure of the region of the human elastin gene, great abundance
0	PT of all residues containing sequence coding sequences.;
0	PL Connect. Tissue Res. 16:197-211(1987).
0	EN [2]

[illegible]

[illegible][illegible]

RESULT	5
ID	014235
ID	Q14235
AC	PRELIMINARY; PRT; 687 AA.
OC	O14235;
DT	01-NOV-1996 (IREMELrel. 01, Created)
ST	01-NOV-1996 (IREMELrel. 01, Last sequence update)
PT	01-OCT-2000 (IREMELrel. 15, Last annotation update)
DE	ELASTIN.
EN	EJLN.
OS	Bono sapiens (Human).
SC	Chordata; Vertebrata; Euteleostomi;
NCBI_taxid	9606;
Mammalia; Eutheria; Primates; Catarrhini, Homidae; Homo.	
NBCL_faxid	5606;
[1]	

Indik Z., Yoon K., Morrow S.D., Cicilia G., Rosenbloom J., Rosenbloom J., Ornstein-Goldstein N. "Structure of the 3' region of the human elastin gene: great abundance of Alu repetitive sequences and few coding sequences"; *J. molec. tissue Res.* 16:197-211 (1987).

[CITE] SEQUENCE FROM N.A.

INDIK Z., YEH H., Ornstein-Goldstein N., Sheppard P., Anderson N., Rosenbloom J., Citronson L., Rosenbloom J., Ornstein-Goldstein N. "Analysis of cloned genomic and complementary DNA"; *Proc. Natl. Acad. Sci. USA* 84:1171-1175 (1987).

[illegible]

Query Match
59.6%; Score 606; DB 4; Length 687;


```
FT NON_TER 1 1
SQ SEQUENCE 679 AA; 57652 MW; EB3C019B39D7618D CRC64;

Query Match
  48 28; Score 490; DB 6; Length 679;
  Best Local Similarity 38 88; Pred. No. 7 3e-24;
  Matches 146; Conservative 12; Mismatches 34; Indels 184; Gaps 16;

  QY 1 PFGVGVGIGIPGVAGV---PGVG---GVPGVG-----GVP----- 30
  DB 116 PFGIGV-LPGVFTGAGVKPKARGGGGAFAGIPGVFGGQGPGLGYPKAPKPAAGV 174
  QY 31 -----GVGISPERQAAAAAKAAKY----- 49
  DB 175 LPFKTKLPFGFGGAGVAGSAGKAGTCTGCTGVTGFGQAAAAAKAAKLGAGGAGVLPGV 234
  QY 50 -----GVGTP-AAAAAKAAKAAQFGL-----VPGVGVPAGV 81
  DB 235 GGAGIPGAPGAIPIGIGGAGVAGPAAAAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 293
  QY 82 VAPGVGAPVAGVGL---APGVGV---APGVGV---APGVGVAPALGP----- 118
  DB 294 V-PGVGV-PGVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 351
  QY 119 -----PEQAAAAAKAAKY----- 132
  DB 352 KFGAGGVGIGIPFGVGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPG 411
  QY 133 -----GVGTPAAAAAKAAKAAQFGLPVGVGVPAGVPGVPGVPGVPGVPGV 180
  DB 412 AIPGVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 471
  QY 181 VVAPGVGVPAGVGA 196
  DB 472 IGLGPGVPGVAGVPA 487

RESULT 8
Q28098 PRELIMINARY; PRT; 707 AA.
AC Q28098
DT 01-NOV-1996 (TRENDEL. 01, Created)
DI 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DE 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DE ELASTIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 17-35 FROM N.A.
RX MEDLINE=85280426; PubMed=2992576;
RA Cicilia G., May K., Ornstein-Goldstein N., Indik Z., Morrow S.,
RA Rosenbloom J.C., Yoon K., Rosenbloom J., Yoon K.;
RT "Structure of the 3' portion of the bovine elastin gene.";
RL Biochemistry 24:3075-3080(1985).
RN [2]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=88028442; PubMed=3665402;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicilia G., Yoon K., Rosenbloom J.;
RT "Sequence variation of bovine elastin mRNA due to alternative
RT splicing.";
RL Biol. Molec. Biol. 7:235-247(1987).
DR ENBL; M19372; AAA30498.1; JOINED.
DR ENBL; M11422; AAA30498.1; JOINED.
DR ENBL; M19365; AAA30498.1; JOINED.
DR ENBL; M19367; AAA30498.1; JOINED.
DR ENBL; M19368; AAA30498.1; JOINED.
DR ENBL; M19369; AAA30498.1; JOINED.
DR ENBL; M19370; AAA30498.1; JOINED.
DR ENBL; M19371; AAA30498.1; JOINED.
```

```
DR ENBL; M22771; AAA30498.1; JOINED.
DR ENBL; M22772; AAA30498.1; JOINED.
DR ENBL; M22773; AAA30498.1; JOINED.
DR ENBL; M22774; AAA30498.1; JOINED.
DR ENBL; M22775; AAA30498.1; JOINED.
DR ENBL; M22988; AAA30498.1; JOINED.
DR ENBL; M23010; AAA30498.1; JOINED.
DR INTERPRO: IPR00104; -.
DR INTERPRO: IPR001459; -.
DR PRINTS: PRO0308; ANTIPEPSEI.
DR PRINTS: PRO0959; MEVGALKINASE.
FT NON_TER 1
SQ SEQUENCE 707 AA; 60346 MW; FDFD559B8B34CE33 CRC64;

Query Match
  48 28; Score 490; DB 6; Length 707;
  Best Local Similarity 38 88; Pred. No. 7 6e-24;
  Matches 146; Conservative 12; Mismatches 34; Indels 184; Gaps 16;

  QY 1 PFGVGVGIGIPGVAGV---PGVG---GVPGVG-----GVP----- 30
  DB 116 PFGIGV-LPGVFTGAGVKPKARGGGGAFAGIPGVFGGQGPGLGYPKAPKPAAGV 174
  QY 31 -----GVGISPERQAAAAAKAAKY----- 49
  DB 175 LPFKTKLPFGFGGAGVAGSAGKAGTCTGCTGVTGFGQAAAAAKAAKLGAGGAGVLPGV 234
  QY 50 -----GVGTP-AAAAAKAAKAAQFGL-----VPGVGVPAGV 81
  DB 235 GGAGIPGAPGAIPIGIGGAGVAGPAAAAAKAAKAAKAAKAAKAAKAAKAAKAAKAA 293
  QY 82 VAPGVGAPVAGVGL---APGVGV---APGVGV---APGVGVAPALGP----- 118
  DB 294 V-PGVGV-PGVGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 351
  QY 119 -----PEQAAAAAKAAKY----- 132
  DB 352 KFGAGGVGIGIPFGVGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPGGPG 411
  QY 133 -----GVGTPAAAAAKAAKAAQFGLPVGVGVPAGVPGVPGVPGVPGVPGV 180
  DB 412 AIPGVGPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGVPGV 471
  QY 181 VVAPGVGVPAGVGA 196
  DB 472 IGLGPGVPGVAGVPA 487

RESULT 9
Q28096 PRELIMINARY; PRT; 666 AA.
AC Q28096
DT 01-NOV-1996 (TRENDEL. 01, Created)
DI 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DE 01-NOV-1996 (TRENDEL. 01, Last sequence update)
DE ELASTIN (FRAGMENT).
OS Bos taurus (Bovine).
OC Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE OF 17-35 FROM N.A.
RX MEDLINE=85280426; PubMed=2992576;
RA Cicilia G., May K., Ornstein-Goldstein N., Indik Z., Morrow S.,
RA Rosenbloom J.C., Yoon K., Rosenbloom J., Yoon K.;
RT "Structure of the 3' portion of the bovine elastin gene.";
RL Biochemistry 24:3075-3080(1985).
RN [2]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=88028442; PubMed=3665402;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicilia G., Yoon K., Rosenbloom J.;
```


[illegible]

Search completed: April 24, 2001, 16:40:27
Job time: 417 sec

Query-Match 25.6%; Score 260; DB 5; Length 988;
Best Local Similarity 36.8%; Pred. No. 2.8e-09;
Matches 75; Conservative 5; Mismatches 112; Indels

RESULT						15
ID	Q28100	PRT:	76 AA.			
AC	Q28100;	PRELIMINARY:				
D	DT Q28100;					
DD	DT 01-NOV-1996 (TRENDEL.) OI, Created)					
DE	DT 01-NOV-1996 (TRENDEL.OI, Last sequence update)					
DI	DT 01-OCT-2000 (TRENDEL.15, Last annotation update)					
DR	EELSTN (FRAGMENT).					
DS	Bos taurus (Bovine).					
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pec-					
OCC	Bovidae; Bovinae; Bos.					
R	NCBI TAXID=9913;					
RP	[1]					
RR	SEQUENCE FROM N.A.					
RA	MEDLINE=6509254; PubMed=6150137;					
RX	Rosenblum J.;					
RL	"Elastic": relation of protein and gene structure to					
RL	Lab. invest. S1:65-62(1)1984).					
RD	EMBL; M1834; RAASG416.1.					
RF	EMBL; M1834; RAASG416.1. JOINED.					
DR	EMBL; M1834; RAASG416.1.					
DN	NOT TER.					
SF	SEQUENCE 76 AA; 6619 MW; B683379DAES7B202 CRG64.					

```
Query Match      24.5%; Score 249.5; DB 6; Length 76;
Best Local Similarity 60.8%; Pred. No. 1.1e-09;
Matches 62; Conservative 2; Mismatches 9; Indels 29; Gaps
```

```
QY 52 GTPAAA-AKAAAKAATGTVPGVGAFGVGGVAGFGVAGFAGVGAVFGVAPG 109
   ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    2 GVPAAKSAALAAAK-AQFLRGLGVGGVAGFGVGGVGGVAGFAGVAGNGI-GFG----- 54
```



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sv model

Run on: April 24, 2001, 16:38:15 ; Search time 115.25 seconds
(without alignments)
1.984 Million cell updates/sec

Title: US-09-340-736-3

Perfect score: 18

Sequence: 1 KAK 4

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Aligned: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database: A_Geneseq_Q401.*

1: /SIDSL/gcgdata/geneseq/AA1980.DAT.*

2: /SIDSL/gcgdata/geneseq/AA1981.DAT.*

3: /SIDSL/gcgdata/geneseq/AA1982.DAT.*

4: /SIDSL/gcgdata/geneseq/AA1983.DAT.*

5: /SIDSL/gcgdata/geneseq/AA1984.DAT.*

6: /SIDSL/gcgdata/geneseq/AA1985.DAT.*

7: /SIDSL/gcgdata/geneseq/AA1986.DAT.*

8: /SIDSL/gcgdata/geneseq/AA1987.DAT.*

9: /SIDSL/gcgdata/geneseq/AA1988.DAT.*

10: /SIDSL/gcgdata/geneseq/AA1989.DAT.*

11: /SIDSL/gcgdata/geneseq/AA1990.DAT.*

12: /SIDSL/gcgdata/geneseq/AA1991.DAT.*

13: /SIDSL/gcgdata/geneseq/AA1992.DAT.*

14: /SIDSL/gcgdata/geneseq/AA1993.DAT.*

15: /SIDSL/gcgdata/geneseq/AA1994.DAT.*

16: /SIDSL/gcgdata/geneseq/AA1995.DAT.*

17: /SIDSL/gcgdata/geneseq/AA1996.DAT.*

18: /SIDSL/gcgdata/geneseq/AA1997.DAT.*

19: /SIDSL/gcgdata/geneseq/AA1998.DAT.*

20: /SIDSL/gcgdata/geneseq/AA1999.DAT.*

21: /SIDSL/gcgdata/geneseq/AA2000.DAT.*

22: /SIDSL/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	8	R91774	Proteinase A subst
2	18	100.0	8	R91774	Proteinase A subst
3	18	100.0	8	B08138	Peptide modulating
4	18	100.0	8	B08138	Peptide modulating
5	18	100.0	9	R79675	Cyclin B/p33(cdc2)
6	18	100.0	9	R79675	Cyclin B/p33(cdc2)
7	18	100.0	10	Y57974	Xenopus histone H1
8	18	100.0	10	W48036	AE101 alanine scan
9	18	100.0	11	R04180	Monomer of repetat
10	18	100.0	11	B25025	SNP detection meth
11	18	100.0	11	Y84932	Peptide recognised
12	18	100.0	12	R88700	Allergen Alt a 12

Lys/Ala diastereom
Nucleic acid conde
Antipathogenic pep
Nucleic acid conde
Nucleic acid conde
Antipathogenic pep
Antipathogenic pep
Antipathogenic pep
Sequence of immuno
Glucagon precursor
Murine MHC class I
Interleukin-1 rece
Pactolus protein a
AAK79 AA38/639 mu
Nucleic acid conde
Antimicrobial pro-
Human AKAP 79 pept
Mutant peptide AKA
BPI derived peptid
BPI protein segmen
BPI.23, BPI domain
BPI.23 for use in
Anti-fungal BPI pe
Bacterial permeabi
Bacterial permeabi
Antiphilic antihi
Recombinant BPI pe
Peptide #7 derived
Peptide #8 derived
Peptide #9 derived
Bactericidal/perme
Xagalin I or II a
Human BPI protein
Histone H1 deri
Peptide BPI 119.

ALIGNMENTS

RESULT 1
R91774
ID R91774 standard; peptide; 6 AA.
XX
XX R91774;
XX
XX
XX 03-SEP-1996 (first entry)
XX
XX Proteinase A substrate peptide #17.
XX
XX Proteinase A, peptide substrate; activity: beer froth: colour;
XX fluorescence; cleavage.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 6 /label= OTHER
XX /note= "p-nitro-L-Phe"
XX
XX JP08067694-A.
XX
XX PD 12-MAR-1996.
XX
XX 21-JUN-1995; 95JP-0154797.
XX
XX 21-JUN-1994; 94JP-0138704.
XX (SUNR) SUNTORY LTD.
XX WPI; 1996-196571/20.
XX Peptide substrate for the determ. of proteinase A activity - useful
XX to determine stability of beer froth

XX Example 2; Page 7; 9pp; Japanese.
PS
CC Peptides p91759-80 are specific examples of a novel generic proteinase A.
CC peptide substrate (R91757) and are modified with a substrate to develop a
CC detectable colour or fluorescence after cleavage of the Phe-Phe bond by
CC proteinase A. The peptides are used to determine proteinase A activity
CC in a sample e.g. beer froth.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 18; DB 17; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAAK 4
DB 1 kaak 4
|||||

RESULT 2
B08137 standard; peptide: 8 AA.
XX B08137;
AC
DT 04-DEC-2000 (first entry)
XX
DE Peptide modulating activity of heparin, and other glycans.
XX
KW Glycoaminoglycan; proteoglycan; heparin modulation; anticoagulant;
KW cell attachment; cell adhesion; vein graft; tumour cell metastasis;
KW cartilage differentiation; wound healing.
XX
OS Synthetic.

Key Location/Qualifiers
FH Key
FT Misc-difference 1.8
FT /note- "this peptide may be repeated an
FT unspecified number of times"
XX
PN W0200045831-AL.
XX
PD 10-AUG-2000.
XX
PF 02-FEB-1999; 99US-0118276.
XX
PR 02-FEB-2000; 2000WO-US02853.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI San Antonio JD, Verrecchio A, Schick BP;
XX
DR WPI; 2000-543446/49.
XX

PT Novel synthetic peptides with high affinity for glycoaminoglycans and
PT proteoglycans, useful for modulating heparin, promoting cell
PT attachment, modulating tumour metastasis and modulating wound healing -
XX
PS Disclosure: Page 23; 76pp; English.

CC The present sequence represents a synthetic peptide which has a high
CC affinity for glycoaminoglycans and proteoglycans. The peptide is useful
CC in methods for modulating heparin or other glycoaminoglycans with
CC anticoagulant activity, promoting cell attachment or adhesion to
CC tumour cell metastasis, modulating cartilage differentiation, targeting
CC drugs to epithelial cell surfaces (or to other cells expressing
CC proteoglycans), modulating enzymes that act on glycoaminoglycan
CC substrates, affinity purification of bioactive sequences of a
CC glycoaminoglycan, modifying endothelial cell pro-coagulant or
CC anti-coagulant functions mediated through glycoaminoglycans, and

CC modulating wound healing. The peptide may also be used for blocking
CC tissue uptake of heparin or other glycoaminoglycans in a mammal to
CC increase heparin half-life in circulation.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 18; DB 21; Length 8;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAAK 4
DB 4 kaak 7
|||||

RESULT 3
B08138
ID B08138 standard; peptide: 8 AA.
XX
AC B08138;
XX
DT 04-DEC-2000 (first entry)
XX
DE Peptide modulating activity of heparin, and other glycans.
XX
KW Glycoaminoglycan; proteoglycan; heparin modulation; anticoagulant;
KW cell attachment; cell adhesion; vein graft; tumour cell metastasis;
KW cartilage differentiation; wound healing.
XX
OS Synthetic.

Key Location/Qualifiers
FH Key
FT Misc-difference 1.8
FT /note- "this peptide may be repeated an
FT unspecified number of times"
XX
PN W0200045831-AL.
XX
PD 10-AUG-2000.
XX
PF 02-FEB-2000; 2000WO-US02853.
XX
PR 02-FEB-1999; 99US-0118276.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI San Antonio JD, Verrecchio A, Schick BP;
XX
DR WPI; 2000-543446/49.
XX

PT Novel synthetic peptides with high affinity for glycoaminoglycans and
PT proteoglycans, useful for modulating heparin, promoting cell
PT attachment, modulating tumour metastasis and modulating wound healing -
XX
PS Disclosure: Page 23; 76pp; English.

CC The present sequence represents a synthetic peptide which has a high
CC affinity for glycoaminoglycans and proteoglycans. The peptide is useful
CC in methods for modulating heparin or other glycoaminoglycans with
CC anticoagulant activity, promoting cell attachment or adhesion to
CC tumour cell metastasis, modulating cartilage differentiation, targeting
CC drugs to epithelial cell surfaces (or to other cells expressing
CC proteoglycans), modulating enzymes that act on glycoaminoglycan
CC substrates, affinity purification of bioactive sequences of a
CC glycoaminoglycan, modifying endothelial cell pro-coagulant or
CC anti-coagulant functions mediated through glycoaminoglycans, and

Query Match 100.0%; Score 18; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAKK 4
 DB 2 kaak 5

RESULT 4
 B08144
 ID B08144 standard; peptide: 8 AA.

XX AC B08144;

XX 04-DEC-2000 (first entry)

XX Peptide modulating activity of heparin, and other glycans.

XX Glycoaminoglycan; proteoglycan; heparin modulation; anticoagulant;
 KW cell attachment; cell adhesion; vein graft; tumour cell metastasis;
 KW cartilage differentiation; wound healing.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference /note= "this peptide may be repeated an
 PT unspecified number of times"

XX WO200045831-A1.

XX 10-AUG-2000.

XX 02-FEB-2000; 2000MO-US02853.

XX 02-FEB-1999; 9905-0118276.

XX (UKJB-) UNIV JEFFERSON THOMAS.

XX San Antonio JD, Verrecchio A, Schick BP;

XX WPI; 2000-543446/49.

XX Novel synthetic peptides with high affinity for glycoaminoglycans and
 proteoglycans, useful for modulating heparin, promoting cell
 attachment, modulating tumour metastasis and modulating wound healing -

XX Disclosure; Page 23; 76pp; English.

XX The present sequence represents a synthetic peptide which has a high
 affinity for glycoaminoglycans and proteoglycans. The peptide is useful
 in methods for modulating heparin or other glycoaminoglycans with
 anticoagulant activity, promoting cell attachment or adhesion to
 natural or synthetic surfaces (especially vein grafts), modulating
 tumour cell metastasis, modulating cartilage differentiation, targeting
 drugs to epithelial cell surfaces (or to other cells expressing
 proteoglycans), modulating enzymes that act on glycoaminoglycan
 substrates, affinity purification of bioactive sequences of a
 glycoaminoglycan, modifying endothelial cell pro-coagulant or
 anticoagulant activity, and modulating glycoaminoglycans and
 modulating wound healing. The peptide may also be used for
 tissue uptake of heparin or other glycoaminoglycans in a mammal to
 increase heparin half-life in circulation.

XX Sequence 8 AA;

Query Match 100.0%; Score 18; DB 21; Length 8;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAKK 4
 DB 4 kaak 7

RESULT 5

XX R79675 standard; peptide: 9 AA.

XX AC R79675;

XX 26-FEB-1996 (first entry)

XX Cyclin B/p33(cdc2) phosphorylation site #2 in Xenopus histone H1.

XX Peptide library; phosphorylation site; protein kinase; substrate;
 KW inhibitor; competitor; cellular response; cell cycle control;
 KW immune response; transcriptional activation; cell development.

XX Synthetic.

XX WO9518823-A2.

XX 13-JUL-1995.

XX 06-JAN-1995; 95WO-US00147.

XX 07-JAN-1994; 94US-0178570.

XX (BETH-) BETH ISRAEL HOSPITAL ASSOC.

XX Cantley LC, Songyang Z;

XX WPI; 1995-255036/33.

XX Determin. of amino acid sequence of protein kinase phosphorylation
 site - by phosphorylation of peptide library and sequencing
 PT phosphopeptide(s) formed, also new substrates and their analogues
 PT for modulating or detecting protein kinase

XX Example 6; Page 34; 131pp; English.

XX An oriented degenerate peptide library of the amino acid formula
 CC R79661 was constructed to isolate the amino acid sequences at the
 CC phosphorylation sites of a protein kinase eg. protein kinase A,
 CC cyclin B/p33(cdc2), src family kinases, etc. peptides which are
 CC phosphorylated are isolated and their amino acid sequences are compared
 CC to known substrate/inhibitor peptide sequences for that protein kinase.
 CC The peptides R79674-88 represent phosphorylation sites for the cell cycle
 CC control kinases cyclin B/p33(cdc2) or cyclin B/p33(CDK2). This peptide
 CC sequence is the second phosphorylation site in the Xenopus histone H1.
 CC The isolated peptides can be used to screen cDNA libraries for effects on the
 CC protein kinase activity, screening cDNA libraries for protein kinase
 CC substrates or modulators, screening cDNA libraries for protein kinase
 CC kinases are involved eg. cell cycle control, immune response,
 CC transcriptional activation or cell development.

XX Sequence 9 AA;

Query Match 100.0%; Score 18; DB 16; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAKK 4
 DB 1 kaak 4

RESULT 6

XX Y57974 standard; peptide: 9 AA.

XX AC Y57974;
 XX DT 23-MAR-2000 (first entry)
 XX DA Xenopus histone H1 peptide SEQ ID NO:28.
 XX KW Protein kinase: phosphorylation site; signal transduction.
 XX OS Xenopus sp.
 XX PN US6004757-A.
 XX PD 21-DEC-1999.
 XX PF 06-JUN-1995; 95US-0369643.
 XX PR 07-JAN-1994; 94US-0178570.
 XX PA (BETH-) BETH ISRAEL HOSPITAL ASSOC.
 XX PI Cantley LC, Songyang Z;
 XX WP: 2000-096301/08.
 XX PT Peptide substrate for a kinase, useful for determining substrate
 PT specificity -
 XX Example 6; Column 31; 59pp; English.
 XX The present invention describes a substrate for 1ck comprising a 9 amino
 CC acid peptide (I), also described is a method of inhibiting kinase
 CC activity of 1ck by contacting it with (I) in vitro. The peptide is
 CC useful for determining substrate specificity of a protein kinase.
 CC Information on the substrate specificity of protein kinases in signal
 CC transduction would provide an insight into signal transduction
 CC mechanisms and could allow for the design of therapeutic agents. The
 CC present sequence represents a peptide used in the exemplification of
 CC the present invention.
 XX Sequence 9 AA;
 SQ
 Query Match 100.0%; Score 18; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAK 4
 DB ||||
 1 kaak 4
 RESULT 7
 ID W48036 standard; peptide; 10 AA.
 XX W48036;
 XX DT 12-JUN-1998 (first entry)
 XX DE AE101 alanine scanning analogue effector compound SEQ ID NO:114.
 XX KW Mammalian T1 key peptide; mammalian invariant chain protein; allergy;
 XX immune response; MHC class II; antigenic; autoimmune disease.
 XX OS Synthetic.
 XX PN Key Location/Qualifiers
 XX FT Modified-site 1
 XX FT Modified-site /note= "Acylated"
 XX FT Modified-site 10
 XX FT Modified-site /note= "Amidated"
 XX FT

PN W09749430-A1.
 XX 31-DEC-1997.
 XX 09-JUN-1997; 97WO-US09993.
 XX 26-JUN-1996; 96US-067605.
 XX (ANTI-) ANTIGEN EXPRESS INC.
 XX Adams S, Humphreys RE, Xu M;
 XX WP: 1998-076917/07.
 XX New mammalian invariant chain protein (Ii) key peptide(s) - used for
 PT modulation of immune response, e.g. for treating malignant, allergic
 PT or autoimmune disease of allograft rejection
 XX Example 2; Page 40; 149pp; English.
 XX The present sequence represents an AE101 analogue effector compound
 CC used in the present invention. The present invention describes
 CC a mammalian invariant chain protein (Ii) key peptide of sequence
 CC LNMKLPKPKVSMR and modifications with the exclusion of peptide
 CC YNMKLPKPKVSMR. MHC class II molecules are synthesised in the
 CC endoplasmic reticulum with their antigenic peptide sites blocked by the
 CC invariant chain protein (Ii). The products and method can be used for
 CC the modulation of an immune response for therapeutic or diagnostic
 CC purposes. The enhancement of immunity can be used in the treatment of
 CC e.g. malignant or allergic disease. The immunosuppression can be used
 CC for the treatment of autoimmune disease, e.g. rheumatoid arthritis,
 CC Crohn's disease, diabetes mellitus, lupus erythematosus, and
 CC psoriasis of allograft rejection.
 XX Sequence 10 AA;
 SQ
 Query Match 100.0%; Score 18; DB 19; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAK 4
 DB ||||
 7 kaak 10
 RESULT 8
 ID R04180 standard; protein; 11 AA.
 XX R04180;
 XX AC R04180;
 XX DT 10-SEP-1990 (first entry)
 XX DE Monomer of repetitive primary structure of a synthetic fibrous
 XX protein.
 XX KW Collagen; elastin; keratin; troponin C; silk; dopa protein;
 XX synthetic skin; cosmetics; bioadhesive.
 XX OS Synthetic.
 XX PN Key Location/Qualifiers
 XX FT Misc-difference 9
 XX FT /label=May be Phe or Tyr.
 XX PN W09003438-A.
 XX DT 05-APR-1990.
 XX PF 06-SEP-1989; 89WC-0003839.
 XX PR 30-SEP-1988; 88US-0251714.

XX (ALLC) ALLIED SIGNAL INC.
 XX Goldberry I, Salerno AJ;
 XX WPI; 1990-132274/17.
 XX New bacterial strains for heterologous gene expression -
 XX contg. elements for initiating activity, retarding proteolysis
 XX and stabilising heterologous genes.
 XX Disclosure; : lpp; English.
 XX The polymer of the peptide may act as an analogue to peptides
 XX contg. a repetitive or quasi-repetitive structure eg. collagen,
 XX elastin, keratin, tropoin C, dopa proteins, silk proteins,
 XX bioadhesive proteins and insect cuticle proteins. These products
 XX may be useful in the manufacture of fibrous products, synthetic
 XX skin and cosmetic additives.
 XX Sequence 11 AA;
 SQ

Query Match 100.0%; Score 18; DB 11; Length 11;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAX 4
 Db 5 kaak 8
 ||||

RESULT 9
 ID B29025 standard; peptide; 11 AA.
 AC B29025;
 XX 26-JAN-2001 (first entry)
 XX SNP detection method labelling antibody peptide SEQ ID NO: 1.
 XX Single nucleotide polymorphism; SNP; antibody; label; rabbit; cancer;
 XX neurodegenerative disease; neurological disease; metabolic disease;
 XX autoimmune disease.
 XX Oryctolagus cuniculus.
 PN WO200056926-A2.
 XX 28-SEP-2000.
 XX 15-MAR-2000; 2000WO-US06950.
 XX 19-MAR-1999; 99DS-0272970.
 XX (VALI-) VALIGENE CORP.
 PA (TREI/) TREICH I.
 XX Treich I, Iris FJM, Pournay J;
 XX WPI; 2000-647157/62.
 XX High-throughput nucleotide polymorphism analysis using peptide-labeled
 XX oligonucleotide probes, useful for analysis of gene expression and
 XX detecting alternatively spliced RNA transcripts -
 XX Disclosure; Page 33; 62pp; English.
 PS The present invention relates to methods for detecting single nucleotide
 XX polymorphisms (SNPs) using peptide-labelled oligonucleotides and antibody
 XX arrays. Oligonucleotide probes with identifiable labels are used in the
 XX ValigeneSN Mutation Screening Peptide-Linked (VMSM-PL) method. These

CC Labels may be peptides or antibodies. The present sequence is an example
 CC of the peptide which may be used as a substrate for the detection of
 CC mutations which are associated with various human diseases,
 CC including cancers, neurological diseases, metabolic diseases,
 CC cardiovascular diseases and autoimmune diseases. In addition, it can be
 CC used to detect and diagnose Lyme disease, tuberculosis and sexually
 CC transmitted diseases.
 XX Sequence 11 AA;
 SQ

Query Match 100.0%; Score 18; DB 21; Length 11;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAX 4
 Db 7 kaak 10
 ||||

RESULT 10
 ID B4932 standard; peptide; 11 AA.
 AC B4932;
 XX 21-AUG-2000 (first entry)
 XX Peptide recognised by rabbit IgG antibodies.
 XX IgG antibody; cDNA library; nucleic acid sorting; gene expression.
 XX Oryctolagus cuniculus.
 PN WO200023622-A1.
 XX 27-APR-2000.
 XX 15-OCT-1999; 99WO-US23906.
 XX 16-OCT-1998; 98US-0174328.
 XX (VALI-) VALIGENE CORP.
 PA Iris PJ, Pournay J;
 XX WPI; 2000-339716/29.
 XX Peptide-Labelled Oligonucleotide methods for manipulating, e.g.
 XX labeling, isolating or screening, populations of nucleic acids involve
 XX the use of distinguishable and identifiable peptide tags linked to
 XX oligonucleotide primers -
 XX Disclosure; Page 40; 61pp; English.
 XX The present sequence represents a peptide which is recognised by rabbit
 XX IgG antibodies. The antibody that recognises the present peptide may
 XX be used in the method of the invention. The specification describes a
 XX method for sorting a mixture of nucleic acids derived from cDNA
 XX libraries. The method comprises labeling DNA from each of the libraries
 XX by polymerase chain reaction (PCR) using distinguishably-labelled
 XX oligonucleotide primers for each library, contacting differently
 XX labelled DNA so hybridisation can occur, and sorting DNA using one or
 XX more molecules, each capable of binding one of the labels. The method
 XX is useful for analysing complex mixtures containing nucleic acids.
 XX Alternatively, the method is useful for sorting populations of nucleic acids.
 XX The method is further useful for monitoring gene expression. The method is also useful for
 XX increasing and supplementing the analytical powers of other techniques
 XX of manipulating complex cDNA populations.
 XX Sequence 11 AA;
 SQ

Query Match 100.0%; Score 18; DB 21; Length 11;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAK 4
 Db 7 kaak 10

RESULT 11

R88700 R88700 standard; Peptide: 12 AA.
 ID R88700;
 XX
 XX R88700;
 XX
 DT 19-AUG-1996 (first entry)
 DE Allergen Alt a 12 T-cell epitope 1.
 DE Alt a 45; Alt a 12; allergen; epitope; immunoglobulin E; IGE;
 KW detection; immunotolerance; energy.
 KW Alternaria alternata.
 XX
 XX AF9402038-A.
 XX
 PD 15-NOV-1995.
 XX
 XX 02-NOV-1994; 94AT-0002038.
 XX
 PR 02-NOV-1994; 94AT-0002038.
 XX
 PA (BIOM-) BIONAY PRODN & HANDELSGES MSH.
 XX
 XX Ratz G, Breitenbach M, Ebner C, Kraft D, Rechenauer E;
 PI Oberkofler H, Prillinger H, Simon B, Unger A;
 XX WPI; 1996-040555/95.
 XX
 PT Recombinant DNA encoding allergens of Alternaria alternata - useful
 PT in diagnosis and treatment of A. alternata allergies
 XX
 PS Claim 1; Page 15; 21pp; German.
 XX
 CC R88700-01 are T-cell epitopes derived from the Alt a 12 allergen protein
 CC (R88622) isolated from Alternaria alternata. Peptide epitopes from Alt a
 CC 12 and Alt a 45 (R88672) are useful as diagnostic reagents, e.g. for in
 CC vitro detection of allergy caused by Alt a 45 and 12 (by reaction with
 CC IgG in serum). They can also detect cellular reaction to the specified
 CC allergen. The peptides can also be used therapeutically to induce immunotolerance
 CC or allergy of T lymphocytes.
 XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 18; DB 17; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAK 4
 Db 7 kaak 10

RESULT 12

W35172
 ID W35172 standard; peptide; 12 AA.
 XX
 XX W35172;
 XX
 DT 14-APR-1998 (first entry)
 XX

DE Lys/Ala diastereomer peptide [D]-A3,4,8,10-K4A8.
 XX
 KW Lys/Ala diastereomer peptide; infection; therapy; excitatory neurotoxin;
 KW honey bee venom; pardaxin cytolytic activity; cancer;
 KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
 KW agricultural, pesticide; cell wall lysis.
 XX
 OS Synthetic.

XX Key Location/Qualifiers
 XX
 FT MISC-difference 3 /note= "D-form residue"
 FT MISC-difference 4 /note= "D-form residue"
 FT MISC-difference 10 /note= "D-form residue"
 FT MISC-difference 12 /note= "D-form residue"
 FT Modified-site 12 /note= "C-terminal amide"
 XX
 PN W09731019-A2.
 XX
 XX 28-AUG-1997.
 XX
 XX 20-FEB-1997; 97WO-IL00066.
 XX
 XX 22-FEB-1996; 96IL-0117233.
 XX
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX
 PI Oren Z, Shai Y;
 XX
 XX WPI; 1997-435086/40.
 DR
 XX Peptide(s) having selective cytolytic activity - against pathogens
 PT and malignant cells, but no haemolytic activity, used for treating
 PT infections and cancer
 XX
 XX Example 4; Page 45; 80pp; English.
 CC
 CC This sequence represents a Lys/Ala diastereomer peptide of the
 CC invention. The peptides of the invention have: (a) cytolytic activity on
 CC pathogenic cells (pathogens and malignant cells not naturally present in
 CC the body); but (b) no haemolytic activity, or such activity only at a
 CC concentration significantly higher than that at which they lyse
 CC pathogens. The peptides, their complexes and mixtures are used to treat
 CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
 CC or cancer, in human and veterinary medicine. Also, they can be used as
 CC preservatives for food, cosmetics and agricultural produce, or as
 CC agricultural pesticides. The absence of haemolytic activity (associated
 CC with lysis) has the advantage of making the peptides more suitable
 CC for use in vivo if any toxic effects and those that include D-as will have
 CC increased resistance to proteolytic degradation. Non-haemolytic,
 CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
 CC activity, allowing selection of agents for particular applications. Since
 CC these random copolymers induce total lysis of bacterial cell walls,
 CC resistance to them is unlikely to develop.
 XX
 SQ Sequence 12 AA;

Query Match 100.0%; Score 18; DB 18; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAK 4
 Db 9 kaak 12

RESULT 13

W08390

ID W05390 standard; peptide; 12 AA.
 AC W08390;
 DT 19-SEP-1997 (first entry)
 DE Nucleic acid condensing peptide consensus sequence III.
 KW Synthetic virus like particle; nucleic acid; condensing peptide;
 KW heteropeptide; polydispersion; transfection; mammalian cell;
 KW gene therapy; self assembly; consensus.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Misc-difference 1 /label= Lys, Arg
 FT Misc-difference 2 /label= Ala, Thr
 FT Misc-difference 10 /label= Pro, Ala, Ser
 FT Misc-difference 12 /label= Lys, Thr, Val
 PN W05641506-A2.
 PD 27-DEC-1996.
 XX 10-JUN-1996; 95W0-GB01396.
 XX 07-JUN-1996; 96US-0860231.
 PR 18-JUN-1995; 95US-0009124.
 PR 21-SEP-1995; 95GB-0013304.
 PR 25-SEP-1995; 95US-0004285.
 PR 19-DEC-1995; 95GB-0025955.
 PR 19-DEC-1995; 95US-0008952.
 PR 12-FEB-1996; 96US-0011331.
 XX (THER-) THEREXSYS LTD.
 XX Craig KK, Cunliffe VP, Thatcher DR, Welsh JR, Wilks PE;
 PI WPI; 1997-05278/06.
 XX Synthetic virus-like particles for gene therapy - comprising
 XX recombinant nucleic acid and nucleic acid condensing peptide(s)
 PS Claim 36; page 123; 193pp; English.
 CC A novel synthetic virus like particle (SVLP), comprises a
 CC recombinant nucleic acid molecule and a non-covalently associated
 CC set of nucleic acid condensing peptides (NACP), e.g. a peptide
 CC comprising the present consensus sequence, where each NACP is a
 CC heteropeptide, and the NACP set has low polydispersion. The SVLP,
 CC which can be used for gene therapy, is self assembling and can
 CC be designed to target a particular cell or tissue type and deliver
 CC the nucleic acid molecule into its chromosomal or extrachromosomal
 CC sequences.
 XX Sequence 12 AA;
 SQ

Query Match 100.0%; Score 18; DB 18; Length 12;
 Best Local Similarity 100.0%; Pred.No.1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 DB 6 kaaK 9
 W69198
 W69198 standard; peptide; 12 AA.
 AC W69198;

RESULT 14
 W62868
 AC W62868 standard; peptide; 12 AA.
 DT 19-MAY-1999 (first entry)
 DE Antipathogenic peptide.
 KW Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
 KW cancer; infection; disinfectant; contact lens wetting solution;
 KW preservative; pesticide; fungicide; bactericide.
 OS Synthetic.
 PN W09837090-Al.
 PD 27-AUG-1998.
 XX 19-FEB-1998; 98W0-IL00081.
 XX 20-FEB-1997; 97W0-IL00066.
 XX (YEDA) YEDA RES & DEV CO LTD.
 XX Oren Z, Shai Y;
 XX NPI; 1998-59464/50.
 XX New non-haemolytic cytolytic agent useful in treating cancer or
 XX infections - is a peptide comprising a moiety which disrupts the
 XX continuity of an alpha-helical structure
 PS Example 4; Page 44; 126pp; English.
 CC The present peptide is used to produce the agents of the invention. The
 CC specification describes a non-haemolytic, cytolytic agent, which is a
 CC peptide, a complex of bundled peptides, a mixture of peptides or a random
 CC peptide copolymer. The agent has a selective cytolytic activity on
 CC pathogenic cells. The agent is selected from a cyclic derivative of a
 CC peptide which has a net positive charge greater than 1, comprises L-amino
 CC acid residues and/or D-amino acid residues and comprises an alpha-helix
 CC breaker moiety, or a peptide (or cyclic derivative of this), which has
 CC a net positive charge greater than 1 and has an amino acid sequence such that
 CC a corresponding amino acid sequence comprising only L-amino acid residues
 CC is not found in nature. The cytolytic agents may be used for treatment of
 CC cancer or for treatment of several diseases caused by pathogens,
 CC including bacterial, fungal, viral, mycoplasma and protozoan infections.
 CC They may be used in both human and veterinary medicine. They may also be
 CC used as disinfectants for destruction of microorganisms, i.e. in the
 CC solutions for wetting contact lenses, as preservatives, e.g., in the
 CC cosmetic and food industries, as pesticides (e.g. fungicides or
 CC bactericides) or for preservation of agricultural products.
 XX Sequence 12 AA;
 SQ

Query Match 100.0%; Score 18; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred.No.1.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 DB 9 kaaK 12
 W69198

RESULT 15
 W69198
 ID W69198 standard; peptide; 12 AA.
 XX W69198;
 AC

XX 09-OCT-1998 (first entry)
DI Nucleic acid condensing peptide #3.
XX
XX Nucleic acid condensing peptide: cell transfection; gene delivery system.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH FT Misc-difference 1 /label= Lys, Arg
FT FT Misc-difference 2 /label= Ala, Val, Thr
FT FT Misc-difference 10 /label= Pro, Ala, Ser
FT FT Misc-difference 12 /label= Ala, Lys, Thr, Val
XX
XX WO9828625-A1.
FN
XX
XX 02-JUL-1998.
FT
XX 23-DEC-1997; 97WO-GB03523.
XX
XX 27-DEC-1996; 96GB-0026992.
XX 23-DEC-1996; 96US-0035908.
XX
XX (COBR-) COBRA THERAPEUTICS LTD.
PA
XX
XX Thatcher DR, Wilks PE;
PI
XX
XX MPI; 1998-377790/32.
DR
XX
XX Transfection of cells with nucleic acid(s) - comprises use of
FT peptide(s) screened for interaction with nucleic acids on sensor
FT chip by surface plasmon resonance
XX
XX Disclosure: Page 23: 75pp; English.
XX
XX This sequence represents a nucleic acid condensing peptide that was
CC identified using the method of the invention. The method is for screening
CC test peptides for the ability to optimally transfect cells with a nucleic
CC acid molecule (NAM), and comprises: (a) detecting a change in the surface
CC plasmon resonance (SPR) of a nucleic acid immobilised on a sensor chip
CC and exposed to a solution of a test peptide, where the change in SPR
CC occurs upon binding of the peptide to and dissociation of the peptide
CC from the immobilised nucleic acid, to permit calculation of the
CC equilibrium constant (K_d) (or apparent dissociation constant (K_d)); and
CC (b) selecting the peptide having K_d with a value of 1x10⁻¹² to 1x10⁻⁶
CC and the value of K_d is 7x10⁻¹². The methods described herein are used for
CC optimising the efficiency of transfection of cells with NAMs. The method
CC of the invention provides a means for identifying and selecting peptides that
CC provide a gene delivery complex that is stable in the bloodstream,
CC targetable to selected tissue types, and capable of efficient transport
CC into the cytoplasm and to the nucleus.
XX
XX Sequence 12 AA;

Query Match 100.0%; Score 18; DB 19; Length 12;
Best Local Similarity 100.0%; Pred.No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 kAAK 4
DB 6 kAAK 9

Search completed: April 24, 2001, 16:38:17
Job time: 417 sec


```

RESULT 2
US-08-911-364-3
; Sequence 3, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION:
; APPLICANT: ROHSTEIN, Aser
; APPLICANT: KEELY, Fred W.
; APPLICANT: ROHSTEIN, Steven J.
; TITLE OF INVENTION: ELONGATING PEPTIDES MODELED ON HUMAN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911.364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023.552
; FILING DATE: 07-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/POCKET NUMBER: 041082/0104
; TELEPHONE: (202) 872-5300
; TELEFAX: (202) 872-5399
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-911-364-3

Query Match 100.0%; Score 18; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
Db 1 KAAK 4

RESULT 3
US-08-178-570-28
; Sequence 28, Application US/08178570
; Patent No. 5532167
; GENERAL INFORMATION:
; APPLICANT: Lewis C. Cantley
; APPLICANT: Songyang, Zhou
; TITLE OF INVENTION: Substrate Specificity of Protein Kinases
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, suite 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA

```

```

; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/178.570
; FILING DATE: JANUARY 7, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A. Jr.
; REFERENCE/POCKET NUMBER: BRL-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
US-08-178-570-28

Query Match 100.0%; Score 18; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
Db 1 KAAK 4

RESULT 4
US-08-369-643-28
; Sequence 28, Application US/08369643A
; Patent No. 6004757
; GENERAL INFORMATION:
; APPLICANT: Cantley, Lewis C.
; APPLICANT: Songyang, Zhou
; TITLE OF INVENTION: Substrate Specificity of Protein Kinases
; FILE REFERENCE: CNS-001CP
; CURRENT APPLICATION NUMBER: US/08/369.643A
; CURRENT FILING DATE: 1995-01-06
; EARLIER FILING DATE: 1994-01-07
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 28
; LENGTH: 9
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: Xenopus histone
; OTHER INFORMATION: HI
US-08-369-643-28

Query Match 100.0%; Score 18; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
Db 1 KAAK 4

RESULT 5
PCT-US95-00147-28
; Sequence 28, Application PCT/US9500147

```

```

GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Substrate Specificity of Protein Kinases
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS: ROCKFELD
STREET: 60 STATE STREET, SUITE 510
CITY: BOSTON
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00147
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,570
FILING DATE: JANUARY 7, 1994
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-004CPPC
TITLE OF INVENTION: SUBSTRATE SPECIFICITY OF PROTEIN KINASES
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
FRAGMENT TYPE: internal
PCT-US95-00147/28

Query Match 100.0%; Score 18; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 1 KAAK 4

RESULT 6
US-08-968-676-114
Sequence 114, Application US/08968676
Patent No. 5919639
GENERAL INFORMATION:
APPLICANT: Humphreys, Robert E
APPLICANT: Xu, Minzhen
CORRESPONDENCE ADDRESS:
NUMBER OF SEQUENCES: 155
ADDRESS: Kevin M. Farrell, P.C.
STREET: P.O. Box 999
CITY: York Harbor
STATE: ME
COUNTRY: USA
ZIP: 03911
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/968,676
FILING DATE:

```

```

CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: RH-9601
TITLE OF INVENTION: HAPTENATED PEPTIDES AND USES THEREOF
TELEPHONE: (207) 363-0558
TELEFAX: (207) 363-0528
INFORMATION FOR SEQ ID NO: 114:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
FRAGMENT TYPE: internal
US-08-908-616-114

Query Match 100.0%; Score 18; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 7 KAAK 10

RESULT 7
PCT-US95-04121-38
Sequence 38, Application PC/TUS9504121
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Haptenated Peptides and Uses Thereof
NUMBER OF SEQUENCES: 62
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04121
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/222,206
FILING DATE: MAY 1, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Vastano, Darlene A.
REGISTRATION NUMBER: 35,279
REFERENCE/DOCKET NUMBER: 079,2PCT
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6010
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
MOLECULE TYPE: linear
TOPOLOGY: linear
FRAGMENT TYPE: internal
PCT-US95-04121-38

Query Match 100.0%; Score 18; DB 5; Length 13;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 4 KAAK 7

```

```

RESULT 8
US-08-503-226B-37
; Sequence 37: Application US/08503226B
; Patent No. 5871945
; GENERAL INFORMATION:
; APPLICANT: Lockerie, Robert Owen, et al.
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
; TITLE OF INVENTION: Anchoring Protein
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,226B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIORITY NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-6300
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-503-226B-37

Query Match 100.0%; Score 18; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAAK 4
Db 6 KAAK 9

RESULT 9
US-08-503-226B-38
; Sequence 38: Application US/08503226B
; Patent No. 5871945
; GENERAL INFORMATION:
; APPLICANT: Lockerie, Robert Owen, et al.
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
; TITLE OF INVENTION: Anchoring Protein
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,226B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIORITY NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-6300
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-503-226B-38

Query Match 100.0%; Score 18; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAAK 4
Db 6 KAAK 9

RESULT 9
US-08-503-226B-38
; Sequence 38: Application US/08503226B
; Patent No. 5871945
; GENERAL INFORMATION:
; APPLICANT: Lockerie, Robert Owen, et al.
; TITLE OF INVENTION: Calcineurin Inhibitory Compounds and
; TITLE OF INVENTION: Anchoring Protein
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,226B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIORITY NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-6300
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-503-226B-38

Query Match 100.0%; Score 18; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAAK 4
Db 6 KAAK 9

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/503,226B
; FILING DATE: 15-MAR-1995
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/404,731
; FILING DATE: 15-MAR-1995
; PRIORITY NUMBER: US 08/344,227
; FILING DATE: 23-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32861
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-6300
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-503-226B-38

Query Match 100.0%; Score 18; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAAK 4
Db 6 KAAK 9

RESULT 10
US-08-860-150-18
; Sequence 18: Application US/08860150B
; Patent No. 5981205
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; TITLE OF INVENTION: Nuclear DBP2-Related (NDR) Kinases
; TITLE OF INVENTION: 2025/ACTUS/US/08/860,150B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,150B
; FILING DATE: 1997-06-19
; EARLIER FILING DATE: 1995-12-20
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 14
; TYPE: PPT
; ORGANISM: Homo sapiens
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(14)
; OTHER INFORMATION: Predicted carboxy terminus of human Ndr
US-08-860-150-18

Query Match 100.0%; Score 18; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 KAK 4
Db 11 KAK 14

RESULT 11
US-09-338-132-18
Sequence 18, Application US/09338132
Patent No. 6040164
APPLICANT: Hemmings, Brian A.
APPLICANT: Millward, Thomas A.
TITLE OF INVENTION: Nuclear DPF2-related (NDR) Kinases
FILE REFERENCE: 4-20265/A/PCT
CURRENT FILING DATE: 1999-06-22
CURRENT APPLICATION NUMBER: US/09/338,132
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05032
EARLIER FILING DATE: 1995-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05032
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 18
LENGTH: 14
TYPE: PPT
ORGANISM: Homo sapiens
NAME/SEQ: PEPTIDE
FEATURE:
OTHER INFORMATION: Predicted carboxy terminus of human Rdr
US-09-338-132-18

Query Match 100.0%; Score 18; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAK 4
Db 11 KAK 14

RESULT 12
US-09-340-736-37
Sequence 37, Application US/08721458B
Patent No. 6107104
GENERAL INFORMATION:
APPLICANT: Lockerbie, Robert Owen, et al.
TITLE OF INVENTION: Calcineurin Inhibitory Compounds
TITLE OF INVENTION: and Anchoring Protein
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,458B
FILING DATE: 15-MAR-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/404,731
FILING DATE: 15-MAR-1995
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/344,227
FILING DATE: 23-NOV-1994
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/503,226
FILING DATE: 17-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33276
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-6300
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-721-458B-37

Query Match 100.0%; Score 18; DB 3; Length 14;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAK 4
Db 6 KAK 9

RESULT 13
US-08-721-458B-38
Sequence 38, Application US/08721458B
Patent No. 6107104
GENERAL INFORMATION:
APPLICANT: Lockerbie, Robert Owen, et al.
TITLE OF INVENTION: Calcineurin Inhibitory Compounds
TITLE OF INVENTION: and Anchoring Protein
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/721,458B
FILING DATE: 15-MAR-1995
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: US 08/404,731
FILING DATE: 15-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/33276
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-6300
INFORMATION FOR SEQ ID NO: 38:

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 14 amino acids
; TYPE: amino acid
; COMPOSITION: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-721-4588-38

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 3; Length 14;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KAAK 4
DB 6 KAAK 9

```

```

RESULT 14
US-08-311-611A-24
; Sequence 24, Application US/08311611A
; Patent No. 5523288

```

```

; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; RECORD TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Shadd, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: "BPI.23"
US-08-311-611A-24

```

```

Query Match
Best Local Similarity 100.0%; Score 18; DB 1; Length 15;

```

```

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KAAK 4
DB 8 KAAK 11

```

```

RESULT 15
US-08-311-611A-115
; Sequence 115, Application US/08311611A
; Patent No. 5523288
; GENERAL INFORMATION:
; APPLICANT: Cohen, Jonathan
; APPLICANT: Kung, Ada H.C.
; APPLICANT: Lambert, Jr., Lewis H.
; TITLE OF INVENTION: Method for Treating Gram-Negative Bacterial
; TITLE OF INVENTION: Infection by Administration of
; TITLE OF INVENTION: Bactericidal/Permeability-Increasing
; NUMBER OF SEQUENCES: 227
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; RECORD TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,611A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/273,401
; FILING DATE: 11-JUL-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/125,651
; FILING DATE: 22-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Shadd, Jeffrey S.
; REGISTRATION NUMBER: 31,879
; REFERENCE/DOCKET NUMBER: 32251
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: "BPI.119"
; NAME/KEY: Modified-site
; LOCATION: 7
; OTHER INFORMATION: /label= Substituted-Ala
; OTHER INFORMATION: /note= "the alanine at position 7 is beta-1-
; OTHER INFORMATION: naphthyl-substituted."
; NAME/KEY: Modified-site
; LOCATION: 10
; OTHER INFORMATION: /label= Substituted-Ala
; OTHER INFORMATION: /note= "the alanine at position 10 is beta-1-
; OTHER INFORMATION: naphthyl-substituted."
US-08-311-611A-115

```

Query Match 100.0%; Score 18; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. NO. 1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
Db 8 KXAK 11

Search completed: April 24, 2001, 16:36:23
Job time: 303 sec

GenCov version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM4 protein - protein search, using sw model

Run on: April 24, 2001, 16:41:54 ; Search time 74.56 Seconds
(without alignments)
3.687 Million cell updates/sec

Title: US-09-340-736-3
Prefix score: 18
Sequence: 1 KAKK 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sched: 198801 segs, 6872235 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 57:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	100.0	22	D47256	Kinetoplast DNA-associated protein p15 - Crithidia fasciculata (fragment)
2	18	100.0	26	A28108	atrial gland peptide D1 - California sea hare (fragment)
3	18	100.0	27	S05454	histone H2A - brin
4	18	100.0	33	S61846	hryf protein - pse
5	18	100.0	35	B45316	sperm-specific pro
6	18	100.0	36	F69800	hypothetical prote
7	18	100.0	38	P41920	annexin-like 40K p
8	18	100.0	39	A45479	GTP-binding regula
9	18	100.0	45	S04941	protamine phi-3.1
10	18	100.0	45	S04941	protamine phi-3.1
11	18	100.0	45	S10545	actin lyase (EC 4
12	18	100.0	45	A34639	ORP overlapping ca
13	18	100.0	45	B43859	ribosomal protein
14	18	100.0	54	S78292	hypothetical prote
15	18	100.0	54	T44359	ribosomal protein
16	18	100.0	58	S78240	ribosomal protein
17	18	100.0	58	F72128	protein translocas
18	18	100.0	61	H75184	hypothetical prote
19	18	100.0	64	D69350	hypothetical prote
20	18	100.0	64	T22415	archaeal histone a
21	18	100.0	69	S66454	prothymosin-like c
22	18	100.0	69	S66454	prothymosin-like c
23	18	100.0	70	H71177	G-ORF-B protein -
24	18	100.0	72	G42512	hypothetical prote
25	18	100.0	72	A81223	hypothetical prote
26	18	100.0	73	F64031	hypothetical prote
27	18	100.0	78	T22171	hypothetical prote
28	18	100.0	78	S06918	DNA-binding protei
29	18	100.0	78	C19434	hypothetical prote

hypothetical prote
adenosine kinase
regulatory protein
ribosomal protein
actobindin - Acan
LSU ribosomal prot
complement C3 - ax
ribosomal protein
hypothetical prote
sperm-specific pro
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
ribosomal protein

ALIGNMENTS

RESULT 1

D47256
Kinetoplast DNA-associated protein p15 - Crithidia fasciculata (fragment)
C: Crithidia fasciculata
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
C:Accession: D47256

R: Xu, C.; Ray, D. S.
Proc. Natl. Acad. Sci. U.S.A. 90, 1786-1789, 1993

A: Title: Isolation of proteins associated with Kinetoplast DNA networks in vivo.

A: Reference number: A47256; MUID: 93189582

A: Accession: D47256

A: Status: Preliminary

A: Molecule type: Protein

A: Note: sequence extracted from NCBI backbone (NCBI:126909)

Query Match 100.0%; Score 18; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KAKK 4

Db 13 KAKK 16

RESULT 2

A28108

atrial gland peptide D1 - California sea hare (fragment)

C: Species: Aplysia californica (California sea hare)

C: Date: 28-Aug-1989 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C: Accession: A28108

R: Hagde, C.; Suter, S. D.; Blankenship, J. E.; Kurosky, A.

A: Title: Proteolytic processing of egg-laying hormone-related precursors in Aplysia.

A: Reference number: A52727; MUID: 88243802

A: Accession: A28108

A: Molecule type: protein

A: Residues: 1-26 <NRG>

Query Match 100.0%; Score 18; DB 2; Length 26;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 KAKK 4

Db 17 KAKK 20

RESULT 3

S05454

histone H2A - brine shrimp (fragment)
 C:Species: Artemia sp. (brine shrimp)
 C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000
 C:Accession: S05454
 C:Superfamily: histone H2A
 Nucleic Acids Res. 17, 6283-6297, 1989
 A:Title: The 5S rRNA-histone repeat in the crustacean Artemia: structure, polymorphism and
 A:Reference number: S05454; MUID:8936657
 A:Accession: S05454
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-27 <CRD>
 A:Cross-references: EMBL:X14816; NID:95695; PDB:1CA3292L.1; PDB:9L335605
 C:Superfamily: histone H2A

Query Match 100.0%; Score 18; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 |||||
 DB 23 KAAK 26

RESULT 4
 hrp protein - Pseudomonas solanacearum
 C:Species: Pseudomonas solanacearum
 C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
 C:Accession: S61846
 R:van Gijsegem, F.; Gough, C.; Zischek, C.; Niqueux, E.; Airlat, M.; Genin, S.; Barberis, M.
 Mol. Microbiol. 15, 1095-1114, 1995
 A:Title: The hrp gene locus of Pseudomonas solanacearum, which controls the production of
 is a complex.
 A:Reference number: S61846; MUID:95349395
 A:Accession: S61846
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-33 <VNA>
 A:Cross-references: EMBL:214056; NID:950397; PDB:95U0398
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1994
 C:Genetics:
 A:Gene: hrpY

Query Match 100.0%; Score 18; DB 2; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 |||||
 DB 26 KAAK 29

RESULT 5
 sperm-specific protamine-like protein - California mussel (fragment)
 N:Alternate names: phil; PL-III
 C:Species: Mytilus californianus (California mussel)
 C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 16-Feb-1997
 C:Accession: B45316
 R:Carlos, S.; Jucifer, L.; Borrell, I.; Hunt, D.F.; Ausio, J.
 J. Mol. Biol. 185, 181-194, 1995
 A:Title: Sequence and characterization of a sperm-specific histone H1-like protein of My
 A:Reference number: A45316; MUID:93106999
 A:Accession: B45316
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-35 <CAP>
 A:Note: sequence extracted from NCBI backbone (NCBIP:121400)
 C:Keywords: DNA binding

Query Match 100.0%; Score 18; DB 2; Length 35;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 |||||
 DB 24 KAAK 27

RESULT 6
 hypothetical protein yfhe - Bacillus subtilis
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
 C:Accession: F69800
 R:Hust, P.; Ogawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.: Ehrlich, S.D.; Emerson, P.F.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Follmer, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hallio, M
 Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Nau
 Y, M.; Ogawa, K.; Ogihara, A.; Odega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, S.; Sadah, Y.; Sato, I.; Scani
 A:Authors: Schaefer, R.; Schaefer, R.; Scorfano, E.; Scorfano, E.; Scorfano, E.; Sc
 T.: Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili
 A:Reference number: A69580; MUID:96044033
 A:Accession: F69800
 A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-36 <KDB>
 A:Cross-references: GH:299108; GB:AL009126; NID:9563055; PDB:1CAB12679.1; PDB:ell828
 A:Experimental source: strain 108
 C:Genetics:
 A:Gene: yfhe

Query Match 100.0%; Score 18; DB 2; Length 36;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 |||||
 DB 31 KAAK 34

RESULT 7
 annexin-like 40K protein - sooty sea hare (fragment)
 N:Alternate names: lipocortin
 C:Species: Aplysia brasiliana (sooty sea hare)
 C:Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 17-Mar-1999
 C:Accession: PH1920
 R:Radu, U.; Nunez-Regueiro, M.; Cook, R.; Kaetzel, M.A.; Yeung, S.C.J.; Eskin, A.
 J. Neurochem. 61, 1236-1245, 1993
 A:Title: Identification of an annexin-like protein and its possible role in the Aplys
 A:Reference number: PH1920; MUID:9389472
 A:Accession: PH1920
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-38 <RAD>
 A:Experimental source: eye
 C:Superfamily: annexin VII; annexin repeat homology
 C:Keywords: calcium binding

Query Match 100.0%; Score 18; DB 2; Length 38;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 13 KXAK 16

RESULT 8
A5479
GTP-binding regulatory protein G1 alpha-3 chain - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
R:Holzman, B.J.; Kinane, T.E.; West, K.; Soper, B.W.; Kargra, H.; Ausiello, D.A.; Ercole, A.; J. Biochem. 182, 569-576, 1989
A:Title: Sequence and characterization of the sperm-specific protein phi3 from Mytilus
A:Reference number: S04941; MUID:89325302
A:Accession: S10544
A:Superfamily: histone H1
A:Residues: 1-45 <NLS>
C:Keywords: DNA binding; nucleus

Query Match 100.0%; Score 18; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 32 KXAK 35

RESULT 11
S10545
Protamine phi-3.3 - California mussel
N:Alternate names: protamine P4-IV, California mussel
C:Species: Mytilus californianus (California mussel)
C:Accession: S10515
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
R:Ausiello, J.; McFarland, R.
A:Title: Sequence and characterization of the sperm-specific protein phi3 from Mytilus
A:Reference number: S04941; MUID:89325302
A:Accession: S10545
A:Molecule type: protein
A:Residues: 1-45 <NLS>
C:Superfamily: histone H1
C:Keywords: DNA binding; nucleus

Query Match 100.0%; Score 18; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 32 KXAK 35

RESULT 12
A34639
pectin lyase (EC 4.2.2.10) - Ervinia carotovora (fragment)
C:Species: Ervinia carotovora
C>Date: 13-Jul-1990 #sequence_revision 13-Jul-1990 #text_change 08-Oct-1999
R:Nishida, T.; Suzuki, T.; Ito, K.; Kanto, Y.; Takai, K.
A:Title: Cloning, expression and characterization of pectin lyase gene from Ervinia carotovora in Esche
A:Reference number: A34639; MUID:90241256
A:Accession: A34639
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-45 <NLS>
A:Cross-references: GB:S35271; NID:g148473; PIDN:AAA24855.1; PID:g551866
C:Keywords: carbon-oxygen lyase

Query Match 100.0%; Score 18; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 16 KXAK 19

RESULT 10
S10544
Protamine phi-3.2 - California mussel
N:Alternate names: protamine PL-IV.2
C:Species: Mytilus californianus (California mussel)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S10544

Query Match 100.0%; Score 18; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 32 KXAK 35

RESULT 9
S04941
Protamine phi-3.1 - California mussel
N:Alternate names: protamine P4-IV.1
C:Species: Mytilus californianus (California mussel)
C:Accession: S04941
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
R:Ausiello, J.; McFarland, R.
A:Title: Sequence and characterization of the sperm-specific protein phi3 from Mytilus
A:Reference number: S04941; MUID:89325302
A:Accession: S04941
A:Molecule type: protein
A:Residues: 1-45 <NLS>
A:Note: 6-Thr, 17-Thr, and 18-Thr was also found
C:Superfamily: histone H1
C:Keywords: DNA binding; nucleus

Query Match 100.0%; Score 18; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 2.4e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 29 KXAK 32

RESULT 10
S10544
Protamine phi-3.2 - California mussel
N:Alternate names: protamine PL-IV.2
C:Species: Mytilus californianus (California mussel)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S10544

Query Match 100.0%; Score 18; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2.7e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KXAK 4
DB 32 KXAK 35

RESULT 10
S10544
Protamine phi-3.2 - California mussel
N:Alternate names: protamine PL-IV.2
C:Species: Mytilus californianus (California mussel)
C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 11-Jan-2000
C:Accession: S10544

RESULT 13

B43859
 ORF overlapping carboxyl terminus of IpaA - Shigella dysenteriae (fragment)
 C:Species: Shigella dysenteriae
 C>Date: 10-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 18-Nov-1994
 C:Accession: B43859
 R:Zao, R.; Palchaudhuri, S.
 Infect. Immun. 60, 1163-1169, 1992
 A:Title: Nucleotide sequence and transcriptional regulation of a positive regulatory gene
 A:Reference number: A43859; MUID:9217561
 A:Contents: CG097, large invasion plasmid
 A:Accession: B43859
 A>Status: Preliminary
 A:Residues: 1-45 <YAO>
 A:Note: sequence extracted from NCBI backbone (NCBI:85367, NCBIP:85368)

Query Match 100.0%; Score 18; DB 2; Length 45;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 ||||
 DB 26 KAAK 29

RESULT 14

S78292
 ribosomal protein L32', chloroplast - Odontella sinensis chloroplast
 C:Species: Chloroplast Odontella sinensis
 C>Date: 17-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
 C:Accession: S78292
 R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.
 Plant Mol. Biol. Rep. 13, 335-342, 1995
 A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis
 A:Reference number: S78298
 A:Accession: S78292
 A:Contents: Protein; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-54 <KOW>
 A:Cross-references: EMBL:267753; NID:g1185127; PIDN:CAA91665.1; PID:g1185182
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C:Genetics:
 A:Gene: rpl32'
 A:Genome: chloroplast
 A:Superfamily: rice chloroplast ribosomal protein L32
 C:Keywords: chloroplast; protein biosynthesis; ribosome

Query Match 100.0%; Score 18; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 ||||
 DB 26 KAAK 29

RESULT 15

T44359
 ribosomal protein rpsB [imported] - Clostridium histolyticum (fragment)
 C:Species: Clostridium histolyticum
 C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 02-Sep-2000
 C:Accession: T44359
 R:Matsumita, O.; Jung, C.M.; Katayama, S.; Minami, J.; Takahashi, Y.; Okabe, A.
 J. Bacteriol. 181, 923-933, 1999
 A:Title: Gene duplication and multiplicity of collagenases in Clostridium histolyticum.
 A:Reference number: 222752; MUID:99121032
 A:Accession: T44359
 A>Status: Preliminary; translated from GB/EMBL/DDJ
 A:Molecule type: DNA

A:Residues: 1-54 <YATP>
 A:Cross-references: EMBL:AB014075; NID:g366863; PIDN:BA334546.1; PID:g3992650
 A:Experimental source: strain JCM 1403
 A:Note: rpsB
 C:Superfamily: rpsB protein

Query Match 100.0%; Score 18; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 ||||
 DB 44 KAAK 47

Search completed: April 24, 2001, 16:41:56
 Job time: 461 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:42:45 ; Search time 44.88 Seconds
(without alignments)
3.053 Million cell updates/sec

Title: US-09-340-736-3

Perfect score: 18

Sequence: 1 KAK 4

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Scheduled: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
1	18 100.0	35	1 PHIL_MITCA	P35422 mytilus cal
2	18 100.0	45	1 PH13_MITCA	P18660 mytilus cal
3	18 100.0	57	1 RK32_ODDSI	P49564 odontella s
4	18 100.0	58	1 RS2L_CHLPN	Q92940 chlamydia p
5	18 100.0	67	1 RHE_PPRMO	Q74092 pyrococcus
6	18 100.0	71	1 RHE_PPRMO	Q74092 pyrococcus
7	18 100.0	71	1 RHE_PPRMO	Q74092 pyrococcus
8	18 100.0	73	1 YP85_BREIN	P43242 haemophilus
9	18 100.0	78	1 PH10_HOLFU	P14309 holochloria
10	18 100.0	82	1 DNAB_STRAL	P96457 streptomyces
11	18 100.0	84	1 ALAA_MOUSE	P97718 mus musculus
12	18 100.0	85	1 KOC1_ECOHI	Q52277 escherichia
13	18 100.0	86	1 RS30_MYCLE	Q31332 mycobacteri
14	18 100.0	88	1 ACTO_ACAN	P18281 acanthamoeb
15	18 100.0	88	1 RGL1_ARCFU	Q28213 archaeglob
16	18 100.0	91	1 PHIL_MITCA	Q04641 mytilus edu
17	18 100.0	92	1 RHE_PPRMO	Q74092 pyrococcus
18	18 100.0	92	1 RHE_PPRMO	Q74092 pyrococcus
19	18 100.0	95	1 RS18_PRCPR	Q92835 rickettsia
20	18 100.0	96	1 Y335_CHLPR	Q84339 chlamydia t
21	18 100.0	96	1 Y612_CHLPR	Q9P358 chlamydia m
22	18 100.0	97	1 RL28_RICPR	Q92848 rickettsia
23	18 100.0	99	1 YORE_PPRMO	P20297 pyrococcus
24	18 100.0	100	1 RLS_SHEEP	P15547 ovine arles
25	18 100.0	103	1 GLOC_RANCA	P15438 rana catesb
26	18 100.0	104	1 HILL_BOVIN	Q02253 bos taurus
27	18 100.0	104	1 HILL_BOVIN	Q02253 bos taurus
28	18 100.0	109	1 SRS2_YEAST	P80597 saccharomyc
29	18 100.0	110	1 GLA2_SCHPO	Q90112 schistosom
30	18 100.0	110	1 RLAL_ACTAL	P49148 alternaria
31	18 100.0	110	1 VFUS_VACCV	P12258 vaccinia vi
32	18 100.0	110	1 VFUS_VARY	P33815 variola vir
33	18 100.0	112	1 HMGD_DRONE	Q05783 drosophila

ALIGNMENTS

RESULT 1
PHIL_MITCA
ID PHIL_MITCA STANDARD; PRT; 35 AA.
AC P35422: 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-JUN-1994 (Rel. 29, Last sequence update)
DE SPERM-SPECIFIC PROTEIN PH1-1 (PL-III) (SPERM-SPECIFIC PROTEIN-LIKE
DE PROTEIN) (FRAGMENT).
OS Mytilus californianus (California mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6549;
RN [1]
RP SEQUENCE.
RC TISSUE-Sperm.
RA MEDLINE=9310699; PubMed=1671995;
RA "Serios S., Guggiaro C., Barakat, J., Hunt D.P., Ausio J.;
RA "The structure and function of a sperm-specific histone H1-like
RA protein of Mytilus californianus.";
EL 7 Biol. Chem. 268:185-194 (1993).
CC -1- FUNCTION: INVOLVED IN NUCLEAR BASIC PROTEIN TRANSITION: HISTONES
CC ARE REPLACED BY SPERMATID SPECIFIC PROTEINS WHICH ARE THEMSELVES
CC REPLACED BY PROTAMINES IN LATE SPERMATIDS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: SPERM.
DR PIR; B45316; B45316.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Non-ter protein.
ET NON-TER Protein.
SQ SEQUENCE 35 AA; 3759 MW; 97459D27A76F4BD1 CRC64;

Query Match 100.0%; Score 18; DB 1; Length 35;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KAK 4

Db 24 KAK 27

RESULT 2
PH13_MITCA
ID PH13_MITCA STANDARD; PRT; 45 AA.
AC P18660: 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DE 01-OCT-1989 (Rel. 12, Last sequence update)
DE SPERM-SPECIFIC PROTEIN PH1-3 (PL-IV)
DE PROTEIN) (FRAGMENT).
OS Mytilus californianus (California mussel).
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
OC Mytiloidea; Mytilidae; Mytilus.
OX NCBI_TaxID=6549;
RN [1]
RP SEQUENCE.

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 Db 30 KAAK 33

RESULT 5
 ID HPHB_PVRHO STANDARD; PRT: 67 AA.
 AC 074092;
 DT 30-MAY-2000 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE PROBABLE ARCHAEAL HISTONE B.
 GN PH5046.
 OS Pyrococcus horikoshii.

CC Archaea; Buryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
 NCBI_TaxID=33953;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=J3;41137; PubMed=9670194;
 RA Yamamoto S., Sekine M., Baba S.-I., Konugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka K., Nakazawa H., Takamiya M., Ohfuku Y.,
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shitaya H., Kikuchi H.;
 RT "Complete sequence and gene organization of the genome of a hyper-
 thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
 RL DNA Res. 5:55-76(1998).
 CC -1- KODENOME FINES AND COMPACT DNA (35 TO 150 BASE PAIRS) TO FORM
 CC -1- KODENOME FINES AND COMPACT DNA (35 TO 150 BASE PAIRS) TO FORM
 CC INCREASES THE RESISTANCE OF DNA TO THERMAL DENATURATION [2]
 CC SIMILARITY)
 CC -1- SIMILARITY: BELONGS TO THE HNF FAMILY OF ARCHAEAL HISTONES.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: AP000007; BAA30815.1; AUL_INIT.
 DR HSP: P48781; 1EFA.
 DR InterPro: IPR000347;
 KW Pfam: PF00808; CBFD_NFYB_HNF; 1.
 KW DNA-binding; Multigene family.
 SQ SEQUENCE 67 AA; 7366 MW; 00A04ACE5F8FE54A CRC64;

Query Match 100.0%; Score 18; DB 1; Length 67;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 Db 24 KAAK 27

RESULT 6
 ID YALL_RHISN STANDARD; PRT: 59 AA.
 AC P55549;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE HYPOTHEICAL 7.3 KDA PROTEIN YALL.
 GN YALL.
 OS Rhizobium sp. (strain NGR234).

CG Plasmid sym pNGR234.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Rhizobium.
 CX NCBI_TaxID=334;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Baloch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes.";
 RL Nature 387:394-401(1997).

CC -1- SIMILARITY: NONE OBVIOUS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: AB000083; AB01761.1;
 DR EMBL: AB000083; AB01761.1;
 SQ SEQUENCE 69 AA; 7275 MW; 234931622798DF CRC64;

Query Match 100.0%; Score 18; DB 1; Length 69;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 Db 25 KAAK 28

RESULT 7
 ID YVGB_VACCC STANDARD; PRT: 72 AA.
 AC P20566;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE HYPOTHEICAL 8.4 KDA PROTEIN.
 OS Vaccinia virus (strain Copenhagen).
 CC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 CC Orthopoxvirus.
 CX NCBI_TaxID=10249;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=91021077; PubMed=2213722;
 RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RT "The complete DNA sequence of vaccinia virus.";
 RL Virology 179:247-266(1990).

CC COMPLETE GENOME.
 RA Goebel S.J., Johnson G.P., Perkus M.E., Davis S.W., Winslow J.P.,
 RA Paoletti E.;
 RL Virology 179:517-563(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL: M35027; AAA48073.1;
 DR EMBL: M35027; AAA48073.1;
 DR PIR: G42512; G42512.
 KW Hypothetical protein
 SQ SEQUENCE 72 AA; 8369 MW; 2E8597B390272B8 CRC64;

```

Query Match      100.0%; Score 18; DB 1; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 10 KAAK 13

RESULT 8
ID YB85_HAEIN STANDARD; PRT: 73 AA.
AC P44212;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE VERY HYPOTHETICAL PROTEIN HIL485 IN MO-LIKE PROPHAGE PLUMO REGION.
GN HIL485
OS Haemophilus influenzae;
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
ON NCBI_TaxID=727;
RP SEQUENCE FROM N.A.
DR EMBL: U22826; AAC23140.1; -
DR STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kervatage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-L., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Fierback F.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Grah C.L., Fleischmann J.L., Rumpman J.L., Saudek D.M., Rumpman J.L.,
RA Venter J.C., McDonald L.A., Small K.V., Fraser C.M., Smith A.O.,
RA "Whole-genome random sequencing and assembly of Haemophilus
RA influenzae RD."
RL Science 269:496-512(1995).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DB EMBL: U22826; AAC23140.1; -
DR TIGR: H1485;
ON Hypothetical protein.
RP SEQUENCE 73 AA; 8834 MW; AC20081B670923FD CRC64;

Query Match      100.0%; Score 18; DB 1; Length 73;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 23 KAAK 26

RESULT 9
ID PHIO_HOLTU STANDARD; PRT: 78 AA.
AC P14309;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE SPERM-SPECIFIC PROTEIN PHI-O.
OS Holothuria tubulosa (sea cucumber).

```

```

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Holothuroidea; Aspidochirotrata; Aspidochirotrata; Holothuriidae;
OC Holothuria.
ON NCBI_TaxID=7685;
RP SEQUENCE FROM N.A.
DR EMBL: M27408; AAC29204.1; -
DR EMBL: X16364; CAA34410.1; -
DR PIR: S06918; S06918.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW Nuclear protein.
SQ SEQUENCE 78 AA; 8680 MW; ABFA270759514F2C CRC64;

Query Match      100.0%; Score 18; DB 1; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 56 KAAK 59

RESULT 10
ID DNAAJ_STRAL STANDARD; PRT: 82 AA.
AC P86457;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CHAPERONE PROTEIN DNAAJ (FRAGMENT).
GN DNAAJ
OS Streptomyces albus G.
OC Actinomycetales; Streptomyces; Streptomyces; Streptomyces.
ON NCBI_TaxID=1962;
RP SEQUENCE FROM N.A.
DR MEDLINE=97157943; PubMed=9004222;
RX STRAIN=J1074;
RA Grandvalet C., Servant P., Mazodier P.;
RT "Disruption of hspR, the repressor gene of the dnaK operon in
RT Streptomyces albus G.,"
AL Mol. Microbiol. 23:77-84(1997).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1759;
EN [1]
RP SEQUENCE FROM N.A.
RA Seeger K., Harris D., Parkhill J., Barrell B.G., Rastandream M.A.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE S20P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not modified
CC or modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: 299125; CAB16149.1;
CC InterPro: IPR002593;
CC Pfam: PF01649; Ribosomal_S20p; 1.
CC Ribosomal protein: rRNA-binding.
CC SEQUENCE 86 AA; 9509 MW; 13818F2730582790 CRC64;
CC -----
Query Match 100.0%; Score 18; DB 1; Length 86;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 4
DB 46 KAAK 49
-----
RESULT 14
ACTO_ACAK CACA STANDARD; PRT; 88 AA.
AD P12291
RA 01-NOV-1990 (Rel. 15, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE ACTOBIINDIN.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
OX NCBI_TaxID=5755;
EN [1]
RP SEQUENCE.
RA Vandekerckhove J., van Damme J., Vancoppenolle K., Bubb M.R.;
RL The covalent structure of Acanthamoeba acrobindingin.;
CC -!- FUNCTION: IS ABLE TO BIND TWO ACTIN MONOMERS AT HIGH
CC CONCENTRATIONS OF G-ACTIN.
CC -!- SUBUNIT: MONOMER.
CC PIR: A36614; A36614.
DR Actin-binding; Repeat; Methylation; Acetylation.
FT MOD_RES 1 1 ACETYLATION.
FT MOD_RES 35 35 METHYLATION (TRP-).
FT MOD_RES 75 75 METHYLATION (TRP-).
FT REPEAT 15 47 NEARLY IDENTICAL.
FT REPEAT 51 84 NEARLY IDENTICAL.
SQ SEQUENCE 88 AA; 9554 MW; E814A5C21603DA5 CRC64;
-----
Query Match 100.0%; Score 18; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 4
DB 84 KAAK 87
-----

```

```

RESULT 15
RL31_ARCFU STANDARD; PRT; 88 AA.
ID RL31_ARCFU
CC 028213;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L31E.
GN RPL31E OR AF2056.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
CC EMBL: AF2056;
CC [1]_TaxID=2234;
RP SEQUENCE FROM N.A.
RA MEDLINE=9845343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides M.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.R., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artach P., Kalne H.P., Sykes S.M.,
RA Beckwith M., Andrea K.P., Bowman C., Fujita C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.H., Smith R.O., Weiss C.R.,
RA Venter J.C.;
RA "the complete genome sequence of the hyperthermophilic, sulphate-
RA reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -!- SIMILARITY: BELONGS TO THE L31E FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is not modified
CC or modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000961; AB89205.1;
CC TIGR: AF2066;
CC InterPro: IPR000054;
CC Pfam: PF01198; Ribosomal_L31e; 1.
CC PROSITE: PS01144; RIBOSOMAL_L31E; 1.
CC Ribosomal protein.
CC SEQUENCE 88 AA; 10488 MW; 04DE178F487BDF54 CRC64;
-----
Query Match 100.0%; Score 18; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 4
DB 29 KAAK 32
-----
Search completed: April 24, 2001, 16:42:47
Job time: 453 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:40:27 : Search time 125.5 Seconds
(without alignments)
3.736 Million cell updates/sec

Title: US-09-340-736-3
Perfect score: 100
Sequence: 1 KAK 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Aligned: 374700 seqs, 11720915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL15:

1: sp-archaea*

2: sp-bacteria*

3: sp-fungi*

4: sp-human*

5: sp-invertebrate*

6: sp-mammal*

7: sp-mouse*

8: sp-organelle*

9: sp-phage*

10: sp-plant*

11: sp-rodent*

12: sp-unclassified*

13: sp-vertebrate*

14: sp-virus*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	18	100.0	11	2	Q9R446	Q9R446 neisseria 9
2	18	100.0	20	5	Q9R446	Q9R446 bos taurus
3	18	100.0	22	5	Q9R446	Q9R446 neisseria m
4	18	100.0	22	5	Q9R446	Q9R446 neisseria m
5	18	100.0	27	5	Q17096	Q17096 artemia sp
6	18	100.0	35	10	Q9SP02	Q9SP02 arabidopsis
7	18	100.0	36	2	Q31573	Q31573 bacillus su
8	18	100.0	37	2	Q85035	Q85035 mycoplasma
9	18	100.0	37	5	Q61191	Q61191 blepharisma
10	18	100.0	38	5	Q61189	Q61189 blepharisma
11	18	100.0	38	5	Q61190	Q61190 blepharisma
12	18	100.0	40	7	Q47517	Q47517 sue scrofa
13	18	100.0	45	2	Q47512	Q47512 sue scrofa
14	18	100.0	54	2	Q92K15	Q92K15 elastridium
15	18	100.0	54	5	Q97314	Q97314 plasmodium
16	18	100.0	58	2	Q950K5	Q950K5 chlamydia p
17	18	100.0	58	11	Q55067	Q55067 meriones un
18	18	100.0	61	1	Q9V2S1	Q9V2S1 pyrococcus
19	18	100.0	64	1	Q29454	Q29454 archaeoglob

20 18 100.0 64 5 Q20588 Q20588 caenorhabdi

21 18 100.0 67 1 Q9VIF5 Q9VIF5 pyrococcus

22 18 100.0 69 6 Q9VIF5 Q9VIF5 pyrococcus

23 18 100.0 69 6 Q9VIF5 Q9VIF5 pyrococcus

24 18 100.0 69 10 Q9S9F3 Q9S9F3 faagopyrum e

25 18 100.0 69 10 Q9S9F3 Q9S9F3 faagopyrum e

26 18 100.0 69 10 Q9S9F3 Q9S9F3 faagopyrum e

27 18 100.0 70 2 Q53012 Q53012 rhizobium e

28 18 100.0 70 3 Q14377 Q14377 schizosacch

29 18 100.0 72 2 Q87036 Q87036 vibrio chol

30 18 100.0 72 2 Q9K106 Q9K106 neisseria m

31 18 100.0 72 2 Q9K106 Q9K106 neisseria m

32 18 100.0 75 2 Q9K106 Q9K106 neisseria m

33 18 100.0 75 2 Q9K106 Q9K106 neisseria m

34 18 100.0 78 5 Q94582 Q94582 holochuria

35 18 100.0 78 5 Q94582 Q94582 holochuria

36 18 100.0 79 9 Q9K086 Q9K086 neisseria m

37 18 100.0 81 4 Q9P1F3 Q9P1F3 homo sapien

38 18 100.0 81 4 Q9P1F3 Q9P1F3 homo sapien

39 18 100.0 81 5 Q18495 Q18495 styela clav

40 18 100.0 81 5 Q18495 Q18495 styela clav

41 18 100.0 81 14 Q92314 Q92314 beet virus

42 18 100.0 82 6 Q97K43 Q97K43 canis famil

43 18 100.0 84 11 P70648 P70648 ratius norv

44 18 100.0 86 2 Q87326 Q87326 helicobacte

ALIGNMENTS

RESULT 1

Q9R446 PRELIMINARY: FRG: 11 AA.

ID Q9R446

DT 01-MAY-2000 (T-EMBLrel. 13, Created)

DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (T-EMBLrel. 13, Last annotation update)

DE CARBAMOYL-PHOSPHATE SYNTHASE SUBUNIT A (FRAGMENT).

GN CARA.

OS Neisseria gonorrhoeae.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

OX NCBI_TaxID=485;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MS11, AND P41090; Francis J.R.;

RA Brinman J.S.; Francis J.R.;

RT Synthesis of carbamoyl-phosphate synthase genes of Neisseria species.

RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF029363; AAC78453.1; -

DR EMBL; AF029362; AAC78452.1; -

FT NON_TER

SQ SEQUENCE 11 AA; 1178 MW; .0CUTAE8E3DD33694 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 11;
Best Local Similarity 100.0%; Pref. No. 88;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAK 4

DB 6 KAK 9

```

RESULT 2
QZIFLQ
ID QZIFLQ PRELIMINARY: PRT; 20 AA.
AC QZIFLQ
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE 25 KDA PROTEIN P25, PEPTIDE FSA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Ruminantia; Bovidae; Bovinae; Bos.
RN [1], TaxID=9913.
RX MEDLINE=91372400; PubMed=1909972;
EX MEDLINE=91372400; PubMed=1909972;
RA Takahashi M., Tomizawa K., Ishiguro K., Sato K., Onori A., Sato S.,
RA Shiratsuchi A., Uchida T., Imahori K.;
RT "A novel brain-specific 25 kDa protein (p25) is phosphorylated by a
RT Ser/Thr-Pro kinase (PTK II) from tau protein kinase fractions.";
FEES Lett. 289:37-43(1991).
SQ SEQUENCE 20 AA; 2032 MW; 4AAB849A6416997 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 17 KAAK 20

RESULT 3
QZIFLQ
ID QZIFLQ PRELIMINARY: PRT; 22 AA.
AC QZIFLQ
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)
DE CARBAMOYL-PHOSPHATE SYNTHASE SUBUNIT A (FRAGMENT).
GN CARA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
RN [1]
RX MEDLINE=5291461; PubMed=7773412;
EX MEDLINE=5291461; PubMed=7773412;
RA Lawson F.S., Billows F.M., Dillon J.A.;
RT "Organization of carbamoyl-phosphate synthase genes in Neisseria
RT gonorrhoeae includes a large, variable intergenic sequence which is
RT also present in other Neisseria species.";
RL Microbiology 141:0-0(0).
RN [2]
RX MEDLINE=1527;
EX MEDLINE=1527;
RA Brinkman F.S.L., Francis F.M., Dillon J.R.;
RT "Complexity of the variable sequence between the carbamoyl-phosphate
RT synthase genes of Neisseria species.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [1]
RX MEDLINE=AR029561; ARC78447.1;
EX MEDLINE=AR029561; ARC78447.1;
RA KOLTER
RN [1]
RX MEDLINE=22 AA; 2427 MW; 3445E9403B5B6AF6 CRC64;
SQ SEQUENCE 22 AA; 2427 MW; 3445E9403B5B6AF6 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 17 KAAK 20

```

```

Db 17 KAAK 20

RESULT 4
QZIFLQ
ID QZIFLQ PRELIMINARY: PRT; 22 AA.
AC QZIFLQ
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
DE KINETOPLAST DNA-ASSOCIATED PROTEIN (FRAGMENT).
OS Crithidia fasciculata.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Crithidia.
RN [1], TaxID=5636;
RX MEDLINE=93189582; PubMed=8446592;
EX MEDLINE=93189582; PubMed=8446592;
RA Xu C., Ray D.S.;
RT "Isolation of proteins associated with kinetoplast DNA networks in
RT vivo.";
FEES Natl. Acad. Sci. U.S.A. 90:1786-1789(1993).
SQ SEQUENCE 22 AA; 1991 MW; 5B1C4941E0BF952 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 13 KAAK 16

RESULT 5
QZIFLQ
ID QZIFLQ PRELIMINARY: PRT; 27 AA.
AC QZIFLQ
DT 01-MAY-1996 (TREMELrel. 01, Created)
DT 01-MAY-1996 (TREMELrel. 01, Last sequence update)
DT 01-MAY-1996 (TREMELrel. 01, Last annotation update)
DE ARTEMIS 5S RNA (PART5H-B), I.
OS Artemia sp. (Brine shrimp).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Branchiopoda; Anostraca;
OC Artemiidae; Artemia.
RN [1]
RX MEDLINE=6662;
EX MEDLINE=6662;
RA STEIN-PAETHOGENETIC POPULATION FROM EURASIA.;
RA Pruces J., Diaz-Guerra M., Gil I., Renart J.;
RT "Polymorphism and variation of the 5S rRNA segment in different
RT populations";
RL Nucleic Acids Res. 17:6283-6297(1989).
RN [1]
RX MEDLINE=CAA32921.1;
EX MEDLINE=CAA32921.1;
RA DNA-binding.
SQ SEQUENCE 27 AA; 2759 MW; 0AE208B586A63B CRC64;

Query Match 100.0%; Score 18; DB 5; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 23 KAAK 26

RESULT 6
QZIFLQ
ID QZIFLQ PRELIMINARY: PRT; 35 AA.
AC QZIFLQ
DT 01-MAY-2000 (TREMELrel. 13, Created)

```

RA	Winters P., Nigst A., Yamamoto H., Yamane K., Yasumoto K., Yeta K.,
RN	Kusuda K., Yoshida H.P., Matsuda T., Yoshikawa H., Danchin A.,
RT	"Whole genome sequence of the gram-positive bacterium Bacillus
RP	subtilis".
RL	Nature 390:249-256(1997).
[2]	
RC	SEQUENCE FROM N.A.
RD	STRAIN:168;
RE	Kunst F., Ogasawara N., Yoshikawa H., Danchin A.:
RF	Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
RG	
RH	SEQUENCE FROM N.A.
RI	MEDLINE=91101547; PubMed=9046165;
RJ	Yamamoto H., Uchiyama S., Sekiguchi J.:
RK	"Cloning and sequencing of a 27.8-kb nucleotide sequence of the 79
RL	degrees-81 degrees region of the Bacillus subtilis genome containing
RM	the spp locus.";
RN	DNA Res. 3:257-262(1996).
RO	EMBL: Z99108; CAB12679.1;
RP	EXML: DB5082; BAA24471.1;
RQ	SEQUENCE 36 AA; 4277 MW; 2EE595192DD65C0 CRC64;
Query Match	100.0%; Score 18; DB 2; Length 36;
Best Local Similarity	100.0%; Pred.No. 2.7e+02;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KA&K 4
DB	
	31 KA&K 34
RESULT	8
ID	OS5035 PRELIMINARY: PRG: 37 AA.
AC	OS5035;
CD	01-NOV-1998 (TEMBREL. 08, Created)
DT	01-NOV-1998 (TEMBREL. 08, Last sequence update)
DI	01-NOV-1998 (TEMBREL. 08, Last annotation update)
DE	HYPOTHETICAL 4.4 KOA PROTEIN.
DS	MYCOPLASMA FILICUTES; Bacillus/Clostridium group; Mollicutes;
OQ	KYCOPLOSMATACEAE; Mycoplasma.
OC	KCBT_TaxID=2098;
[1]	
RC	SEQUENCE FROM N.A.
RD	STRAIN:PG21;
RE	MEDLINE=93553399; PubMed=9687401;
RF	Reber C., Charron A., Rove J.K., Reber C.; Renaudin J.:
RG	"Partial cloning and characterization of the topoisomerase IV parC and
RH	parE genes of Mycoplasma hominis".
RI	Antimicrob. Agents Chemother. 42:2024-2031(1998).
RL	EMBL: AF036961; AAC33843.1;
RM	Hypothetical protein.
RN	SEQUENCE 37 AA; 4369 MW; 8661E5716143A9B4 CRC64;
Query Match	100.0%; Score 18; DB 2; Length 37;
Best Local Similarity	100.0%; Pred.No. 2.9e+02;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KA&K 4
DB	
	12 KA&K 15
RESULT	9
ID	OS61101 PRELIMINARY: PRG: 37 AA.
AC	OS61101;
CD	01-AUG-1998 (TEMBREL. 07, Created)
DT	01-AUG-1998 (TEMBREL. 07, Last sequence update)
DI	01-NOV-1998 (TEMBREL. 08, Last annotation update)
DE	HYPOTHETICAL 4.4 KOA PROTEIN.
OQ	KYCOPLOSMATACEAE; Mycoplasma.
OC	KCBT_TaxID=2098;
[1]	
RC	SEQUENCE FROM N.A.
RD	STRAIN:PG21;
RE	MEDLINE=93553399; PubMed=9687401;
RF	Reber C., Charron A., Rove J.K., Reber C.; Renaudin J.:
RG	"Partial cloning and characterization of the topoisomerase IV parC and
RH	parE genes of Mycoplasma hominis".
RI	Antimicrob. Agents Chemother. 42:2024-2031(1998).
RL	EMBL: AF036961; AAC33843.1;
RM	Hypothetical protein.
RN	SEQUENCE 37 AA; 4369 MW; 8661E5716143A9B4 CRC64;
Query Match	100.0%; Score 18; DB 2; Length 37;
Best Local Similarity	100.0%; Pred.No. 2.9e+02;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KA&K 4
DB	
	12 KA&K 15
RESULT	9
ID	OS61101 PRELIMINARY: PRG: 37 AA.
AC	OS61101;
CD	01-AUG-1998 (TEMBREL. 07, Created)
DT	01-AUG-1998 (TEMBREL. 07, Last sequence update)
DI	01-NOV-1998 (TEMBREL. 08, Last annotation update)
DE	HYPOTHETICAL 4.4 KOA PROTEIN.
OQ	KYCOPLOSMATACEAE; Mycoplasma.
OC	KCBT_TaxID=2098;
[1]	
RC	SEQUENCE FROM N.A.
RD	STRAIN:PG21;
RE	MEDLINE=93553399; PubMed=9687401;
RF	Reber C., Charron A., Rove J.K., Reber C.; Renaudin J.:
RG	"Partial cloning and characterization of the topoisomerase IV parC and
RH	parE genes of Mycoplasma hominis".
RI	Antimicrob. Agents Chemother. 42:2024-2031(1998).
RL	EMBL: AF036961; AAC33843.1;
RM	Hypothetical protein.
RN	SEQUENCE 37 AA; 4369 MW; 8661E5716143A9B4 CRC64;
Query Match	100.0%; Score 18; DB 2; Length 37;
Best Local Similarity	100.0%; Pred.No. 2.9e+02;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KA&K 4
DB	
	12 KA&K 15
RESULT	9
ID	OS61101 PRELIMINARY: PRG: 37 AA.
AC	OS61101;
CD	01-AUG-1998 (TEMBREL. 07, Created)
DT	01-AUG-1998 (TEMBREL. 07, Last sequence update)
DI	01-NOV-1998 (TEMBREL. 08, Last annotation update)
DE	HYPOTHETICAL 4.4 KOA PROTEIN.
OQ	KYCOPLOSMATACEAE; Mycoplasma.
OC	KCBT_TaxID=2098;
[1]	
RC	SEQUENCE FROM N.A.
RD	STRAIN:PG21;
RE	MEDLINE=93553399; PubMed=9687401;
RF	Reber C., Charron A., Rove J.K., Reber C.; Renaudin J.:
RG	"Partial cloning and characterization of the topoisomerase IV parC and
RH	parE genes of Mycoplasma hominis".
RI	Antimicrob. Agents Chemother. 42:2024-2031(1998).
RL	EMBL: AF036961; AAC33843.1;
RM	Hypothetical protein.
RN	SEQUENCE 37 AA; 4369 MW; 8661E5716143A9B4 CRC64;
Query Match	100.0%; Score 18; DB 2; Length 37;
Best Local Similarity	100.0%; Pred.No. 2.9e+02;
Matches	4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 KA&K 4
DB	
	12 KA&K 15
RESULT	9
ID	OS61101 PRELIMINARY: PRG: 37 AA.
AC	OS61

```

DE HISTONE H3 (FRAGMENT).
GN H3-1.
OS Blepharisma undulans.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Heterotrichida;
CC Blepharisma.
OY NCBI_TaxID=74723;
RN NON_TER 37 38
RP SEQUENCE FROM N.A.
RA MEDLINE=98154410; PubMed=9493359;
RX Bernhard D., Schlegel M.;
RT "Evolution of histone H4 and H3 genes in different ciliate lineages.";
RL J. Mol. Evol. 45:344-354(1998).
DR EMBL; AJ004691; CAA06033.1; -.
FT NON_TER 37 38
SQ SEQUENCE 37 AA; 3884 MW; B7723EFCF9C5F2 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 25 KAAK 28

RESULT 10
ID Q61190 PRELIMINARY; PRT; 38 AA.
AC Q61189;
DT 01-AUG-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
DE HISTONE H3 (FRAGMENT).
GN H3-3.
OS Blepharisma undulans.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Heterotrichida;
CC Blepharisma.
OY NCBI_TaxID=74723;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=98154410; PubMed=9493359;
RX Bernhard D., Schlegel M.;
RT "Evolution of histone H4 and H3 genes in different ciliate lineages.";
RL J. Mol. Evol. 45:344-354(1998).
DR EMBL; AJ004691; CAA06043.1; -.
FT NON_TER 37 38
SQ SEQUENCE 38 AA; 3914 MW; E86D5AFF9CF92C6 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 25 KAAK 28

RESULT 11
ID Q61190 PRELIMINARY; PRT; 38 AA.
AC Q61189;
DT 01-AUG-1998 (TRENBLREL. 07, Created)
DT 01-AUG-1998 (TRENBLREL. 07, Last sequence update)
DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
DE HISTONE H3 (FRAGMENT).
GN H3-2.
OS Blepharisma undulans.
OC Eukaryota; Alveolata; Ciliophora; Spirotrichea; Heterotrichida;
CC Blepharisma.
OY NCBI_TaxID=74723;
RN (1)

```

```

RP SEQUENCE FROM N.A.
RA MEDLINE=98154410; PubMed=9493359;
RX Bernhard D., Schlegel M.;
RT "Evolution of histone H4 and H3 genes in different ciliate lineages.";
RL J. Mol. Evol. 45:344-354(1998).
DR EMBL; AJ004690; CAA06041.1; -.
FT NON_TER 38 39
SQ SEQUENCE 38 AA; 3916 MW; E86D5AFF9CF72C6 CRC64;

Query Match 100.0%; Score 18; DB 5; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 25 KAAK 28

RESULT 12
ID Q29047 PRELIMINARY; PRT; 39 AA.
AC Q29047;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
OS Sus scrofa DOMESTICA SUBUNIT (FRAGMENT).
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
CX NCBI_TaxID=9823;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=93179393; PubMed=8440688;
RX Holtzman E.J., Kinane T.B., West K., Soper B.W., Karga H.,
  Ausiello D.A., Roccolani L.;
RT "Transcriptional regulation of G-protein alpha i subunit genes in LLC-
  PK1 cells and characterization of the porcine G alpha i-3 gene
  promoter.";
RL J. Biol. Chem. 268:3964-3975(1993).
DR EMBL; L07317; AAA31049.1; -.
DR HSP; P10824; LG32.
DR INTERPRO; IPR001019; -.
DR PFAM; PF00503; G-alpha; 1.
KW GTP-binding.
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4333 MW; B029D6305AFDE028 CRC64;

Query Match 100.0%; Score 18; DB 6; Length 39;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
DB 29 KAAK 32

RESULT 13
ID Q47512 PRELIMINARY; PRT; 45 AA.
AC Q47512;
DT 01-NOV-1996 (TRENBLREL. 01, Created)
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)
DT 01-NOV-1998 (TRENBLREL. 08, Last annotation update)
DE PECTIN LYASE (PML) (FRAGMENT).
OS Erwinia carotovora.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Pectobacterium.
OX NCBI_TaxID=554;
RN (1)
RP SEQUENCE FROM N.A.

```

RX MEDLINE=90241256; PubMed=2185758;
 RA Wshida T, Suzuki T, Ito K, Kasio Y, Izaki K.;
 RT "Cloning and expression of pectin lyase gene from *Erwinia carotovora*
 AT in *Escherichia coli*.";
 RL Biochem. Biophys. Res. Commun. 166:801-808(1990).
 KW EMBL; M35271; AAA24855.1; -.
 NW Lyase.
 FT NON_TER 45 45
 SQ SEQUENCE 45 AA; 4483 MW; 431674C06CD5B0D7 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 45;
 Best Local Similarity 100.0%; Pred. No. 3.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 DB 16 KAAK 19
 |||||

RESULT 14
 ID Q97314 PRELIMINARY; PRT; 54 AA.
 AC Q97314;
 DT 01-MAY-1999 (TREMELrel. 10, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMELrel. 10, Last annotation update)
 DE RUVB PROTEIN (FRAGMENT).
 GN RUVB.
 OS Clostridium histolyticum.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
 CC Clostridium.
 OA NCBI_TaxID=1498;
 RN 10
 RS SEQUENCE FROM N.A.
 RC STRAIN=JCM 1403;
 RA Matsushita O.;
 RT "Clostridium histolyticum collagenase.";
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB014075; BAA34546.1; -.
 FT NON_TER 54 54
 SQ SEQUENCE 54 AA; 6312 MW; 492B3EA948933292 CRC64;

Query Match 100.0%; Score 18; DB 2; Length 54;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 DB 44 KAAK 47
 |||||

RESULT 15
 ID Q97314 PRELIMINARY; PRT; 54 AA.
 AC Q97314;
 DT 01-MAY-1999 (TREMELrel. 10, Created)
 DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
 DT 01-NOV-1999 (TREMELrel. 12, Last annotation update)
 DE PFC1025C PROTEIN.
 GN PFC1025C.
 OS Plasmodium falciparum.
 CC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OA NCBI_TaxID=5833;
 RN 10
 RS SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RA Oliver K., Rowman S., Harris D., Lawson D., Quail M., Barrell B.;
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL034559; CAB39063.1; -.
 SQ SEQUENCE 54 AA; 6065 MW; E24DBCEB3DE2492B CRC64;

Query Match 100.0%; Score 18; DB 5; Length 54;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 4
 DB 14 KAAK 17
 |||||

Search completed: April 24, 2001, 16:40:30
 Job time: 420 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:38:17 : Search time 115.25 seconds
(without alignments)
2.480 Million cell updates/sec

Title: US-09-340-736-4
Sequence: 1 Kbaak 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

390729 segs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Genescd_0401.*
1: /SIDS1/gcgdata/genescd/genescp/AA1980.DAT.*
2: /SIDS1/gcgdata/genescd/genescp/AA1981.DAT.*
3: /SIDS1/gcgdata/genescd/genescp/AA1982.DAT.*
4: /SIDS1/gcgdata/genescd/genescp/AA1983.DAT.*
5: /SIDS1/gcgdata/genescd/genescp/AA1984.DAT.*
6: /SIDS1/gcgdata/genescd/genescp/AA1985.DAT.*
7: /SIDS1/gcgdata/genescd/genescp/AA1986.DAT.*
8: /SIDS1/gcgdata/genescd/genescp/AA1987.DAT.*
9: /SIDS1/gcgdata/genescd/genescp/AA1988.DAT.*
10: /SIDS1/gcgdata/genescd/genescp/AA1989.DAT.*
11: /SIDS1/gcgdata/genescd/genescp/AA1990.DAT.*
12: /SIDS1/gcgdata/genescd/genescp/AA1991.DAT.*
13: /SIDS1/gcgdata/genescd/genescp/AA1992.DAT.*
14: /SIDS1/gcgdata/genescd/genescp/AA1993.DAT.*
15: /SIDS1/gcgdata/genescd/genescp/AA1994.DAT.*
16: /SIDS1/gcgdata/genescd/genescp/AA1995.DAT.*
17: /SIDS1/gcgdata/genescd/genescp/AA1996.DAT.*
18: /SIDS1/gcgdata/genescd/genescp/AA1997.DAT.*
19: /SIDS1/gcgdata/genescd/genescp/AA1998.DAT.*
20: /SIDS1/gcgdata/genescd/genescp/AA1999.DAT.*
21: /SIDS1/gcgdata/genescd/genescp/AA2000.DAT.*
22: /SIDS1/gcgdata/genescd/genescp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	100.0	7	22	B37101 Poly(Lys)-Poly(Ala)
2	22	100.0	9	11	R04181 Monomer of repeat
3	22	100.0	20	10	W53595 Peptide structure
4	22	100.0	1	10	W53595 Peptide structure
5	22	100.0	12	18	W51772 Lys/Ala diastereom
6	22	100.0	12	19	W52888 Antipathogenic pep
7	22	100.0	12	21	B17423 Antipathogenic pep
8	22	100.0	13	20	W54055 Histone H1 isoform
9	22	100.0	13	21	W54055 Human histone H1.5
10	22	100.0	15	21	W57363 Human histone H1.5
11	22	100.0	15	21	W57361 Human histone H1.5

12	22	100.0	16	17	W07279
13	22	100.0	16	14	B37235
14	22	100.0	19	21	B76735
15	22	100.0	21	16	B76397
16	22	100.0	23	17	W01634
17	22	100.0	25	21	B18770
18	22	100.0	27	21	B18735
19	22	100.0	29	21	B18736
20	22	100.0	36	21	B18737
21	22	100.0	39	19	B45754
22	22	100.0	45	20	B34063
23	22	100.0	45	20	B34063
24	22	100.0	55	21	B57367
25	22	100.0	56	21	B57367
26	22	100.0	56	21	B57367
27	22	100.0	56	21	B57367
28	22	100.0	59	21	B57367
29	22	100.0	60	21	B57367
30	22	100.0	61	19	B76701
31	22	100.0	63	21	B57273
32	22	100.0	66	21	B57274
33	22	100.0	67	20	B57274
34	22	100.0	77	21	B57275
35	22	100.0	86	21	B57275
36	22	100.0	91	21	B57275
37	22	100.0	105	21	B57275
38	22	100.0	109	21	B57275
39	22	100.0	113	21	B57275
40	22	100.0	113	21	B57275
41	22	100.0	116	20	B57275
42	22	100.0	115	21	B57275
43	22	100.0	119	21	B57275
44	22	100.0	123	20	B57275
45	22	100.0	123	20	B57275

ALIGNMENTS

RESULT 1
ID B37101 standard; peptide; 7 AA.
XX B37101:
XX
XX
XX 28-MAR-2001 (first entry)
XX
XX Poly(Lys)-poly(Ala) motif R2 for protein purification method.
XX
XX Production; purification; hybrid polypeptide; affinity chromatography;
XX sulphated polysaccharide; binding motif; antigen; vaccine.
XX Synthetic.
XX W0200068398-A1.
XX
XX 16-NOV-2000.
XX
XX 11-MAY-2003; 2000WO-FR01282.
XX
XX 11-MAY-1999; 99PR-0006031.
XX
XX (INSP) INST PASTEUR LILLE.
XX (INRM) INST NAT SANTE & RECH MEDICALE.
XX
XX Menozzi F, Lochet C, Pêche K;
XX
XX WPI: 2001-007400/01.
XX
XX Purifying hybrid proteins by affinity chromatography, useful for
XX preparing vaccinating antigens, by incorporating motifs with affinity
XX for sulfated polysaccharide -

Amphiphilic anti
Anti-Histone H1
Oligopeptide-based
Heparin-binding pe
Bacterial permeabi
N-dansylated pepti
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Amino acid sequenc
Human histone H1.5
Human histone H1.5
Zea mays protein f
Copolymer molecule
Copolymer molecule
Cone snail alpha-c
Staphylococcus aur
Zea mays protein f
Zea mays protein f
Hoplymer molecule
Hoplymer molecule
Copolymer molecule
Copolymer molecule
Arabidopsis thalia
Human secreted pro
Copolymer molecule
Arabidopsis thalia
Arabidopsis thalia
Histone H1 isoform
Human histone H1.5
Human histone H1.5
Arabidopsis thalia
Arabidopsis thalia
Protein involved i

PS Claim 4; Page 44; 61pp; French.
 CC The invention relates to a method of producing and purifying a hybrid
 CC polypeptide (I) by affinity chromatography on a support carrying at least
 CC sulphated polysaccharides (SP). (I) consists of a protein (II) and at least
 CC one motif (R1) of formula XIXY12Y3 and 0-5 motifs (R2) of formula
 CC XIX2Y12Y3X34, where XI-X4 = K or R; and Y1-Y3 = any amino acid, but
 CC preferably A, L, I, or V. The total number of R1 and R2 is chosen
 CC according to the affinity required between (I) and SP, provided that if
 CC no R2 are present there must be at least two R1. Optionally purified (I)
 CC is treated with a serine protease to remove R1 and R2, leaving (II). This
 CC peptide sequence represents an example of an R2 motif used in the hybrid
 CC polypeptide. The method comprises inserting the coding sequence for (I)
 CC into a cloning vector, expressing (I) in host cells, and purifying (I) from
 CC host cells then expressed (I) is contacted with immobilized SP around
 CC where (I) includes a vaccinating antigen (Ag), optionally with
 CC proteolytic removal of R1 and R2 for recovery of Ag.

XX Sequence 7 AA;
 XX Query Match 100.0%; Score 22; DB 22; Length 7;
 XX Best Local Similarity 100.0%; Pred. No. 3.2e+05; Indels 0; Gaps 0;
 XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAAAK 5
 DB 2 kaaak 6

RESULT 2
 ID R04181
 XX R04181 standard; protein: 9 AA.
 XX R04181;
 XX 10-SEP-1990 (first entry)
 XX Monomer of repetitive primary structure of a synthetic fibrous
 XX protein.
 XX Collagen; elastin; keratin; troponin C; silk; dopa protein;
 XX synthetic skin; cosmetics; bioadhesive.
 XX Synthetic.

XX W09003438-A.
 XX 05-APR-1990.
 XX 06-SEP-1989; 89WO-0003839.
 XX 30-SEP-1988; 88US-0251714.
 XX (ALLC) ALLIED SIGNAL INC.
 XX Goldberg I., Salerno AJ;
 XX WFI; 1990-132274/17.

XX New bacterial strains for heterologous gene expression.
 XX containing proteolytic activity, retarding proteolysis
 XX and stabilising heterologous genes.
 XX Disclosure; 1pp; English.

XX The polymer of the peptide may act as an analogue to peptides
 XX contg. a repetitive or quasi-repetitive structure eg. collagen,
 XX elastin, keratin, troponin C, dopa proteins, silk proteins,
 XX bioadhesive proteins and insect cuticle proteins. These products
 XX may be useful in the manufacture of fibrous products, synthetic
 XX skin and cosmetic additives.

XX Sequence 9 AA;
 SQ Query Match 100.0%; Score 22; DB 11; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAAAK 5
 DB 3 kaaak 7

RESULT 3
 ID W95595
 XX W95595 standard; peptide: 9 AA.
 XX W95595;
 XX 31-JAR-1999 (first entry)
 XX Peptide structure depicting a peptidoglycan (mucopeptide) complex.
 XX Mucopeptide; MPl; lipopolysaccharide; LPS; inhibitory; septic shock;
 XX arthritis; Group A mucopeptide; inflammatory disease; rheumatoid;
 XX Crohn's disease; psoriasis; peptidoglycan.

XX Synthetic.

XX Key Location/Qualifiers
 FT Modified-site 1 /note= "attached to a sugar complex"
 FT Modified-site 2 /note= "D-form residue; Glu-NH2"
 FT Modified-site 3 /note= "attached to D-Ala which is further connected
 FT to another D-Ala (both residues not shown in
 FT the sequence)"
 FT Misc-difference 6 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT Modified-site 9 /note= "Ala-(MurNac-GlcNac)"

XX W09857657-A1.
 XX 23-DEC-1998.
 XX 17-JUN-1998; 98WO-US12647.
 XX 16-JUN-1998; 98US-0097838.
 XX 18-JUN-1997; 97US-0050087.
 XX 18-JUN-1997; 97US-0877988.
 XX (UIPO) UNIV ROCKEFELLER.

XX Blake MS, Zabriskie JB;
 XX WPI; 1999-080932/07.

XX Identifying peptides with affinity to anti-Group A mucopeptide
 XX antibodies - useful to treat septic shock and experimental arthritis
 XX Disclosure; Fig 3; 47pp; English.

XX The invention relates to a new method of identifying an inhibitory
 XX peptide. The method comprises contacting a selected peptide able to bind
 XX a Group A mucopeptide (MPl) and a lipopolysaccharide (LPS) molecule
 XX with: (a) MPl; and (b) a mucopeptide antibody (Ab1) and selecting a
 XX peptide which inhibits binding of MPl with Ab1. The inhibitory peptide
 XX obtained by the method is used to treat septic shock in mammals. Ab1 is
 XX used to prevent or treat arthritis in animals. The selected peptide is

used in a method for identifying antigenic peptides that are used to prevent or ameliorate experimental arthritis induced in an animal. In its broadest application the antibodies raised against group A mucopolysaccharide can be used for the treatment of inflammatory diseases such as septic shock, and rheumatoid arthritis, Crohn's disease, psoriasis, and for the identification of specific peptides that can be used in such treatment. The present sequence represents the structure of a peptidoglycan (mucopolysaccharide) complex.

Sequence 9 AA;

Query Match 100.0%; Score 22; DB 20; Length 9;
Best Local Similarity 100.0%; Pred. No. 3,2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
|||||
3 KAAAK 7

RESULT 4
R08089
ID R08089 standard; protein; 11 AA.

AC R08089;
DF 01-MAR-1991 (first entry)
DE Antifreeze segment f3 encoded by SS8-1.
KW Synthetic antifreeze polypeptide: cryopreservation;
KW core repetitive sequence.
OS Synthetic.
PN W09013571-A.
XX 15-NOV-1990.

XX 10-MAY-1990; 90MO-US02626.

PR 10-MAY-1989; 89DS-0350481.

PR 10-MAY-1990; 90US-0507715.

XX (DNAP-) DNA PLANT TECH COR.

XX Warren GJ, Mueller GM, McKown RL, Dunsmuir P;

XX WPI: 1990-361428/48.

XX New anti-freeze polypeptides - used for cryo-protection of e.g.
PT foods, medically used biologics, plant prods. or plants during
PT growth.

XX Disclosure: Fig 4; 11pp; English.

XX Synthetic anti-freeze polypeptides (saf) comprise a specifically
CC cleavable site and a region containing at least two "core" segments
CC such as the liner given here. The saf's suppress ice crystal growth
CC by binding to the growing crystal face and blocking sites for
CC further crystal growth. They can be used to maximise retention of
CC important properties of organic materials through freezing and
CC thawing processes. The basic design of the polypeptides is based
CC on known antifreeze polypeptides from insects and fish. e.g. the
CC winter flounder. Q06552, Q06556, Q06558, Q06560-Q06572, R08070-6,
CC R08077-8, R08080-R08093.

XX Sequence 11 AA;

Query Match 100.0%; Score 22; DB 11; Length 11;

Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAAK 5
|||||
Db 7 KAAAK 11

RESULT 5

W35172 standard; peptide; 12 AA.

AC W35172;

DT 14-APR-1998 (first entry)

DE Lys/Ala diastereomer peptide [D]-a3,4,8,10-K4A8.

XX Lys/Ala diastereomer peptide; infection; therapy; excitatory neurotoxin;
KW Honey bee venom; pardaxin; cytolytic activity; cancer;
KW non-haemolytic; preservative; agricultural produce; bacterial cell lysis;
KW agricultural pesticide; cell wall lysis.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 10 /note= "D-form residue"

FT Modified-site 12 /note= "D-form residue"

FT /note= "C-terminal amide"

XX W09731019-A2.

XX 28-AUG-1997.

XX 20-FEB-1997; 97MO-IL000066.

XX 22-FEB-1996; 96IL-0117223.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Oren Z, Shai Y;

XX WPI: 1997-435088/40.

XX Peptide(s) having selective cytolytic activity - against pathogens
PT and malignant cells, but no haemolytic activity, used for treating
PT infections and cancer

XX Example 4; Page 45; 80pp; English.

XX This sequence represents a Lys/Ala diastereomer peptide of the
CC invention. The peptides of the invention have: (a) cytolytic activity on
CC pathogenic cells (pathogens and malignant cells not naturally present in
CC the body); but (b) no haemolytic activity, or such activity only at a
CC concentration significantly higher than that at which they lyse
CC pathogens. The peptides, their complexes and mixtures are used to treat
CC infections (caused by bacteria, fungi, protozoa, mycoplasma or viruses)
CC or cancer, in human and veterinary medicine. Also, they can be used as
CC preservatives for food, cosmetics and agricultural produce, or as
CC agricultural pesticides. The absence of haemolytic activity (associated
CC with the lack of natural structure) means that these peptides
CC have no adverse effects and that they include D-as wall have
CC increased resistance to proteolytic degradation. Non-haemolytic.
CC cytotoxic random copolymers of pardaxin, each has a specific spectrum of
CC activity, allowing selection of agents for particular applications. Since
CC these random copolymers induce total lysis of bacterial cell walls,

CC resistance to them is unlikely to develop.

SQ Sequence 12 AA;

Query Match 100.0%; Score 22; DB 18; Length 12;
Best Local Similarity 100.0%; Pred. No. 23;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5

DB 1 KAAAK 5

RESULT 6

ID W82888 standard; peptide; 12 AA.

AC W82888;

DT 19-MAY-1999 (first entry)

XX Antipathogenic peptide.

XX Non-haemolytic; cytolytic; selective cytolytic activity; pathogen;
cancer; infection; disinfectant; contact lens wetting solution;
preservative; pesticide; fungicide; bactericide.

OS Synthetic.

PN W9837090-A1.

PD 27-AUG-1998.

XX 19-FEB-1998; 96WO-IL00081.

XX 20-FEB-1997; 97WO-IL00066.

XX (YEDA) YEDA RES & DEV CO LTD.

XX Oren Z, Shai Y;

XX WPI; 1998-594464/50.

XX New non-haemolytic cytolytic agent useful in treating cancer or
infections - is a peptide comprising a moiety which disrupts the
continuity of an alpha-helical structure

XX Example 4; Page 44; 126pp; English.

XX The present peptide is used to produce the agents of the invention. The
specification describes a non-haemolytic, cytolytic agent, which is a
peptide, a complex of bundled peptides, a mixture of peptides or a random
peptide copolymer. The agent has a selective cytolytic activity on a
pathogenic cells. The agent is selected from a cyclic derivative of a
peptide which has a net positive charge greater than 1, comprises L-amino
acid residues and/or D-amino acid residues and comprises an alpha-helix
breaker moiety, or a peptide (or cyclic derivative of this) which
(comprises L-amino acid residues and D-amino acid residues, has a net
positive charge greater than 1 and has an amino acid sequence such that
a corresponding amino acid sequence comprising only L-amino acid residues
is not found in nature. The cytolytic agents may be used for treatment of
cancer, or for treatment of several diseases caused by pathogens,
including bacterial, viral, fungal and protozoan infections.
They may be used in both human and veterinary medicine. They may also be
used as disinfectants for destruction of microorganisms, i.e. in
solutions for wetting contact lenses, as preservatives, e.g. in the
cosmetic and food industries, as pesticides (e.g. fungicides or
bactericides) or for preservation of agricultural products.

XX Sequence 12 AA;

Query Match 100.0%; Score 22; DB 21; Length 12;

Query Match 100.0%; Score 22; DB 19; Length 12;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5

DB 1 KAAAK 5

RESULT 7

ID BL7423 standard; peptide; 12 AA.

AC BL7423;

DT 31-OCT-2000 (first entry)

XX Antipathogenic peptide sequence SEQ ID NO:527.

XX Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
vascular endothelial growth factor; matrix metalloproteinase;
asthma; thrombosis; pharmaceutical.

OS Synthetic.

PN WC2000024782-A2.

PD 04-MAY-2000.

XX 25-OCT-1999; 99WO-0525044.

XX 23-OCT-1998; 98US-0105371.

XX 22-OCT-1999; 99US-0428082.

XX (ANGE-) ANGEN INC.

XX Feige U, Liu C, Cheetham J, Boone TC;

XX WPI; 2000-350702/30.

XX Novel composition of matter comprising an Fc domain and
pharmacologically active peptides, useful for treating cancer and
autoimmune diseases -

XX Claim 39; Page 381; 608pp; English.

XX The present invention describes composition of matter (I) comprising an
Fc domain pharmacologically active peptides and linkers, where (I) is:
(X1)a-P1-(X2)b, where: P1 = an Fc domain; X1 and X2 = are each
independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
-(L1)c-P1-(L2)d-P2-(L3)e-P3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
where P1, P2, P3, and P4 = are each independently sequences of
pharmacologically active peptides; L1, L2, L3, and L4 = are each
independently linkers; and a, b, c, d, e, and f = are each independently
0 or 1, provided that at least 1 of a and b is 1. The composition can
have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
activities. DNAs, vectors and host cells from the present invention can
be used for producing pharmaceutical compositions. The compositions can
be used for treating cancer, asthma, thrombosis, or autoimmune diseases.
The use of an Fc domain (rather than a Fab domain) can provide a longer
half-life of incorporated functions such as Fc receptor binding, protein A
binding, and fixation of complement. The present invention also includes
sequences A69526 and B18003 represent nucleotide and amino acid sequences
used in the exemplification of the present invention.

XX Sequence 12 AA;

Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 1 KAAK 5

RESULT 8
ID Y34065
AC Y34065;
DT 23-NOV-1999 (first entry)
DE Histone H1 isoform H1.5 pANCA-reactive fragment (residues 172-184).
KW Ulcerative colitis; histone; H1-like antigen; porin antigen; human;
KW Bacteroides antigen; inflammatory bowel disease; IBD; pANCA; diagnosis;
KW perinuclear anti-neutrophil cytoplasmic antibody; isoform.
XX Homo sapiens.
XX WO9945955-A1.
XX 16-SEP-1999.
XX 12-MAR-1999; 99MO-US05492.
XX 12-MAR-1998; 98US-0041889.
XX (REGC) UNIV CALIFORNIA.
XX Braun J, Cohavy O;
XX WPI, 1999-551215/46.
XX Use of histone H1, porin or Bacteroides antigens as targets for the
PT diagnosis, prevention and treatment of ulcerative colitis -
XX Example 4; Page 124; 134pp; English.
XX The invention provides a method for the diagnosis, prevention and
XX treatment of ulcerative colitis (UC) using a histone H1-like antigen, a
XX porin antigen or a bacteroides antigen as a target antigen. The novel
XX method of diagnosing UC in a subject suspected of having inflammatory
XX bowel disease (IBD) comprises: (1) obtaining a sample from the subject;
XX (2) contacting the sample with a histone H1-like antigen, or perinuclear
XX anti-neutrophil cytoplasmic antibody (pANCA)-reactive fragment, to form a
XX complex of the histone H1-like antigen, or the pANCA-reactive fragment,
XX and antibody to the histone H1-like antigen; and (3) detecting the
XX presence or absence of the complex. Where the presence of the complex
XX indicates that the subject has UC, the pANCA-reactive histone H1-like
XX antigen, or the Bacteroides antigen, is used for the diagnosis, prevention
XX and treatment of UC. The methods can also be used
XX for identifying agents useful for treating UC. The present sequence
XX represents a human histone H1 isoform H1.5 pANCA-reactive fragment.
XX Sequence 13 AA:

Query Match 100.0%; Score 22; DB 20; Length 13;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 7 KAAK 11

RESULT 9
Y57363

ID Y57363 standard; peptide; 13 AA.
XX Y57363;
XX 13-JUN-2000 (first entry)
DE Human histone H1.5 pANCA-reactive fragment (residues 172-184).
KW Ulcerative colitis; inflammatory bowel disease; porin antigen;
KW pANCA; perinuclear anti-neutrophil cytoplasmic antibody; human;
KW histone H1; isoform; NANUC-2.
XX Homo sapiens.
XX US6033864-A.
XX 07-MAR-2000.
XX 12-MAR-1998; 98US-0041889.
XX 12-APR-1996; 96US-0057846.
XX 11-APR-1997; 97US-0837058.
XX (REGC) UNIV CALIFORNIA.
XX Cohavy O, Braun J;
XX WPI, 2000-255695/22.
XX Diagnosing ulcerative colitis or susceptibility, by detecting complex
XX formation between microbial porin antigen and perinuclear
XX anti-neutrophil cytoplasmic autoantibodies -
XX Example 4; Columns 67-68; 49pp; English.
XX The invention provides a method for diagnosing ulcerative colitis in a
XX subject suspected of having inflammatory bowel disease. The method
XX comprises reacting a patient sample with a porin antigen that is
XX immunologically reactive with pANCA (perinuclear anti-neutrophil
XX cytoplasmic antibodies) and detecting formation of a Ag-pANCA complex
XX as indicative of ulcerative colitis. The method is used to diagnose
XX ulcerative colitis or susceptibility to it. Sequences 17362-367
XX represent pANCA-reactive fragments, derived from human histone H1.5.
XX Sequence 13 AA:

Query Match 100.0%; Score 22; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 7 KAAK 11

RESULT 10
Y34063
ID Y34063 standard; peptide; 15 AA.
XX Y34063;
XX 23-NOV-1999 (first entry)
DE Histone H1 isoform H1.5 pANCA-reactive fragment (residues 170-184).
KW Ulcerative colitis; histone; H1-like antigen; porin antigen; human;
KW Bacteroides antigen; inflammatory bowel disease; IBD; pANCA; diagnosis;
KW perinuclear anti-neutrophil cytoplasmic antibody; isoform.
XX Homo sapiens.
XX WO9945955-A1.
XX PN

than 5 and 4 respectively, correspond to greater antimicrobial activity.
 CC Effective antimicrobial peptides are discriminated from ineffective
 CC peptides by a vectorial analysis on dimensions corresponding to charge,
 CC hydrophile/lipophile balance, hydrophobic moment and amphiphilicity of
 CC the peptides, where effective peptides fall into the region which is
 CC anti-clockwise of the amphiphilicity dimension and clockwise of the
 CC charge dimension. Effective peptides are further discriminated from
 CC ineffective peptides by means of an equation relating certain physical
 CC properties of the peptides to their biological activity against specific
 CC microorganisms. The peptides can be used in antimicrobial, pref.
 CC disinfection of surfaces, spoilage prevention, preservation or other
 CC hygiene processes.
 XX
 XX Sequence 16 AA:

SEQUENCE 16 AA:

Query Match 100.0%; Score 22; DB 17; Length 16;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 DB 2 KAAAK 6

RESULT 13
 R32756
 ID R32756 standard; peptide: 18 AA.

XX
 XX R32756;

XX
 XX 21-JUN-1993 (first entry)

DE Murine epsilon RMP internal sequence.

XX Epsilon receptor modulating protein; IgE; CD23; class switching;
 KW allergy inflammation; partial sequence.

XX
 XX Mus musculus.

XX
 XX WO9302696-A.

XX
 XX 18-FEB-1993.

XX
 XX 06-AUG-1992; 92WO-US06553.

XX
 XX 07-AUG-1991; 91US-0741671.

XX
 XX (MED-1) MEDICAL BIOLOGY INST.

XX
 XX Katz DH, Matsushita S;

XX
 XX WPI; 1993-076175/09.

XX Immunologically active lig. for e.g. research into allergy etc.
 PT contg. epsilon-receptor modifying protein of specified
 PT characteristics, used opt. with specified
 PT immuno-globin-E-suppressive protein etc.

XX
 XX disclosure; Page 34; 75pp; English.

XX The peptide represents an internal sequence of the 17 kD epsilon
 CC receptor modulating protein (RMP). The protein may be used to
 CC modulate the immune response. The epsilon RMP conjugated with
 CC II-4 can initiate class switching to IgE. This conjugate is also
 CC also useful for treating allergic reactions; e.g. inflammation and
 CC in immunological research and in immunological therapy on humans and
 CC other animals. See also R32755.

XX
 XX Sequence 18 AA:

Query Match 100.0%; Score 22; DB 14; Length 18;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 DB 13 KAAAK 17

RESULT 14

R32755
 ID R32755 standard; peptide: 19 AA.

XX
 XX R32755;

XX
 XX 21-JUN-1993 (first entry)

DE Oligopeptide based on internal sequence of murine epsilon RMP.

XX Epsilon receptor modulating protein; IgE; CD23; class switching;
 KW allergy inflammation; partial sequence.

XX
 XX Synthetic.

XX
 XX Key Location/Qualifiers

FT Modified-site 19 /note= "amidated"

XX
 XX WO9302696-A.

XX
 XX 18-FEB-1993.

XX
 XX 06-AUG-1992; 92WO-US06553.

XX
 XX 07-AUG-1991; 91US-0741671.

XX
 XX (MED-1) MEDICAL BIOLOGY INST.

XX
 XX Katz DH, Matsushita S;

XX
 XX WPI; 1993-076175/09.

XX Immunologically active lig. for e.g. research into allergy etc.
 PT contg. epsilon-receptor modifying protein of specified
 PT characteristics, used opt. with specified
 PT immuno-globin-E-suppressive protein etc.

XX
 XX disclosure; Page 13; 75pp; English.

XX The synthetic sequence is a 19 amino acid sequence with cysteine
 CC inserted into its N-terminus whose sequence was based on the
 CC internal amino acid sequence of purified epsilon receptor modulating
 CC protein. The synthetic oligopeptide was conjugated to keyhole
 CC limpet haemocyanin and used to immunise 8 week old female rats.
 CC The conjugate was used to generate B cell hybridomas secreting PO.
 CC A B cell hybridoma clone secreted antibody that reacted with the
 CC immunogen and with native epsilon RMP.

XX
 XX See also R32756.

XX
 XX Sequence 19 AA:

Query Match 100.0%; Score 22; DB 14; Length 19;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 DB 14 KAAAK 18

RESULT 15

Y87855

ID Y87835 standard; peptide; 19 AA.
AC Y87835;
XX
XX
XX 01-SEP-2000 (first entry)
XX
XX Heparin binding peptide Lys helix #1.
XX
XX Heparin binding peptide; antagonist; cardiovascular; coagulant;
KW bleeding wound; vascular anastomoses; leaking prosthetic vascular graft;
KW protamine substitute; treatment.
XX
XX Synthetic.
OS
XX
XX EF999219-A2.
PN
XX
XX 10-MAY-2000.
FD
XX
XX 01-OCT-1999; 99EP-0119514.
PZ
XX
XX 06-OCT-1998; 98US-0166930.
PR
XX (COMM-) COMMONWEALTH BIOTECHNOLOGIES INC.
PI
Harris RB, Sobel M;
PI
WP; 2000-306006/27.
DR
XX
XX New heparin binding molecules, useful for reducing heparin content in a
PT mammal by reducing the anticoagulant effects of heparin -
XX
XX Example 1; Page 8; 38pp; English.
PS
XX
XX This invention describes novel heparin binding molecules (I). The
CC molecules (I) are useful as heparin antagonist drugs for cardiovascular
CC application and specifically neutralize heparin's conventional
CC anticoagulant properties. (I) are also useful for counteracting actions
CC of heparin locally e.g. in bleeding wounds, vascular anastomoses or
CC leaking prosthetic vascular grafts. (I) is also useful combined in a
CC pharmaceutical composition with insulin, as a substitute for protamine
CC for use in treating diabetics. The heparin binding molecules (I)
CC specifically neutralize heparin's conventional anticoagulant properties
CC the proliferative vascular response to injury. (I) are short duration of
CC intravenous drugs to be used in elective or emergency situations which
CC can safely and specifically neutralize heparin's proliferative response
CC to injury. This sequence represents a heparin-binding peptide described
CC in the method of the invention.
XX
XX Sequence 19 AA;
SQ

Query Match 100.0%; Score 22; DB 21; Length 19;
Best Local Similarity 100.0%; Fred. No. 36;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 K2AAY 5
DB 15 kaak 19

Search completed: April 24, 2001, 16:38:18
Job time: 418 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:35:23 ; Search time 62.39 Seconds
(without alignments)
1.940 Million cell updates/sec

Title: US-09-340-736-4
Perfect score: 22
Sequence: 1 KAAAK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Method: 185757 segs, 1921087 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA*
1: /cnp2_5/prodata/2/aaa/aa_COMB.pep*
2: /cnp2_5/prodata/2/aaa/aa_COMB.pep*
3: /cnp2_5/prodata/2/aaa/aa_COMB.pep*
4: /cnp2_5/prodata/2/aaa/aa_COMB.pep*
5: /cnp2_5/prodata/2/aaa/aa_COMB.pep*
6: /cnp2_5/prodata/2/aaa/aa_COMB.pep*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	100.0	5	2	US-08-911-364-4
2	22	100.0	13	3	US-09-041-889-37
3	22	100.0	15	3	US-09-041-889-35
4	22	100.0	18	5	PCT-US92-06553-2
5	22	100.0	19	2	US-08-660-592-5
6	22	100.0	19	4	US-09-186-930A-5
7	22	100.0	19	5	PCT-US92-06553-1
8	22	100.0	21	1	US-08-473-344-57
9	22	100.0	21	1	US-08-473-344-58
10	22	100.0	21	1	US-08-473-344-59
11	22	100.0	23	5	PCT-US92-065224-1
12	22	100.0	55	3	US-09-041-889-41
13	22	100.0	60	1	US-08-346-849-16
14	22	100.0	60	2	US-08-293-284A-16
15	22	100.0	116	3	US-09-041-889-38
16	22	100.0	158	3	US-09-041-889-40
17	22	100.0	174	3	US-08-911-853-19
18	22	100.0	199	2	US-08-800-264A-11
19	22	100.0	199	2	US-09-048-628-11
20	22	100.0	199	3	US-09-048-628-11
21	22	100.0	199	3	US-09-209-605-11
22	22	100.0	201	2	US-08-911-364-2
23	22	100.0	214	3	US-09-041-889-27
24	22	100.0	218	3	US-09-041-889-4
25	22	100.0	218	3	US-08-837-038-4
26	22	100.0	219	2	US-08-557-309B-54
27	22	100.0	222	3	US-09-041-889-3

Sequence 3, Appli
Sequence 201, App
Sequence 32, Appli
Sequence 1, Appli
Sequence 51, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 8, Appli
Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-911-364-4
; Sequence 4, Application US/08911364
; Patent No. 5563106
; GENERAL INFORMATION:
; APPLICANT: KOTHEIN, Aser
; APPLICANT: KOTHEIN, Steven J.
; TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
; NUMBER OF SEQUENCES: ELKIN AND OTHER FIBROUS PROTEINS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,552
; FILING DATE: 07-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; REPRESENTATION NUMBER: 29,768
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; Molecule type: peptide
US-08-911-364-4

Query Match 100.0%; Score 22; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 Db 1 KAAK 5

RESULT 2
 US-09-041-889-37
 ; Sequence 37, Application US/09041889
 ; Patent No. 6033864
 ; GENERAL INFORMATION:
 ; APPLICANT: Brahm, Jonathan
 ; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
 ; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof. Using
 ; TITLE OF INVENTION: Microbial UC pANCA antigens
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell & Flores LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/041-889
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY INFORMATION DATA:
 ; APPLICATION NUMBER: US 08/837,058
 ; FILING DATE: 11-APR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-PM 3006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 37:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; APPLICATION NUMBER: US/09/041-889
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/837,058
 ; APPLICATION DATE: 11-APR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-PM 3006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 37:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-041-889-37

Query Match 100.0%; Score 22; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 Db 7 KAAK 11

RESULT 3
 US-09-041-889-35
 ; Sequence 35, Application US/09041889
 ; Patent No. 6033864
 ; GENERAL INFORMATION:
 ; APPLICANT: Brahm, Jonathan
 ; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
 ; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof. Using
 ; TITLE OF INVENTION: Microbial UC pANCA antigens
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell & Flores LLP
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/041-889
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIORITY INFORMATION DATA:
 ; APPLICATION NUMBER: US 08/837,058
 ; FILING DATE: 11-APR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-PM 3006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 35:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; APPLICATION NUMBER: US/09/041-889
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/837,058
 ; APPLICATION DATE: 11-APR-1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Campbell, Cathryn A.
 ; REGISTRATION NUMBER: 31,815
 ; REFERENCE/DOCKET NUMBER: P-PM 3006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 35:
 ; LENGTH: 13 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-09-041-889-35

Query Match 100.0%; Score 22; DB 3; Length 13;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 Db 7 KAAK 11

RESULT 4
 PCT-US92-06553-2
 ; Sequence 2, Application PCT/US9206553
 ; GENERAL INFORMATION:
 ; APPLICANT: Katz M.D., David H
 ; APPLICANT: Matsushita M.D., Sho
 ; TITLE OF INVENTION: T-Cell Membrane Protein for B Cell IG
 ; TITLE OF INVENTION: Class Switching
 ; TITLE OF INVENTION: Class Switching
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Grant L. Hubbard
 ; STREET: 2326 North Seventh Street
 ; CITY: Phoenix
 ; STATE: AZ
 ; COUNTRY: USA
 ; ZIP: 85006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/06553
 ; FILING DATE: 19920806
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE: 07-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hubbard, Grant L.

QY 1 KAAK 5
 Db 9 KAAK 13

RESULT 4
 PCT-US92-06553-2
 ; Sequence 2, Application PCT/US9206553
 ; GENERAL INFORMATION:
 ; APPLICANT: Katz M.D., David H
 ; APPLICANT: Matsushita M.D., Sho
 ; TITLE OF INVENTION: T-Cell Membrane Protein for B Cell IG
 ; TITLE OF INVENTION: Class Switching
 ; TITLE OF INVENTION: Class Switching
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Grant L. Hubbard
 ; STREET: 2326 North Seventh Street
 ; CITY: Phoenix
 ; STATE: AZ
 ; COUNTRY: USA
 ; ZIP: 85006
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/06553
 ; FILING DATE: 19920806
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE: 07-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hubbard, Grant L.

REGISTRATION NUMBER: 24193
 REFERENCE/DOCKET NUMBER: M01032PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (602) 263-8782
 TELEFAX: (602) 265-5984
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 18 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE: internal
 CELL TYPE: T-cell
 CELL LINE: NBI-1.15
 PUBLICATION INFORMATION:
 DOCUMENT NUMBER: US 07/741671
 FILING DATE: 07-AUG-1991
 US92-06553-2

Query Match 100.0%; Score 22; DB 5; Length 18;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 DB 13 KAAAK 17

RESULT 5
 US-08-660-592-5

Sequence 5; 697153
 GENERAL INFORMATION:
 APPLICANT: HARRIS, Robert B.
 APPLICANT: SOBEL, Michael
 TITLE OF INVENTION: NOVEL HEPARIN BINDING PEPTIDES
 NUMBER OF SEQUENCES: 11
 CORRESPONDENCE ADDRESS:
 STREET: P.O. Box 1404
 CITY: Alexandria
 STATE: Virginia
 COUNTRY: United States
 ZIP: 22313-1404
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA: US/08/660,592
 APPLICATION NUMBER: US/08/660,592
 CLASSIFICATION: 519-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: McGowan, Malcolm K.
 REGISTRATION NUMBER: 39,300
 REFERENCE/DOCKET NUMBER: 006338-001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 836-6620
 TELEFAX: (703) 836-2021
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19
 TYPE: amino acid acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-660-592-5

Query Match 100.0%; Score 22; DB 2; Length 19;

Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KAAAK 5
 DB 15 KAAAK 19

RESULT 6
 US-09-166-930A-5

Sequence 5; Application US/09166930A
 GENERAL INFORMATION:
 APPLICANT: HARRIS, Robert B.
 APPLICANT: SOBEL, Michael
 TITLE OF INVENTION: NOVEL HEPARIN BINDING PEPTIDES
 FILE REFERENCE: 006338-006
 CURRENT APPLICATION NUMBER: US/09/166,930A
 CURRENT FILING DATE: 1998-10-06
 PRIOR APPLICATION NUMBER: US 08/660,592
 PRIOR FILING DATE: 1998-06-11
 NUMBER OF SEQ ID NOS: 6
 SEQ ID NOS: 1-6
 LENGTH: 19
 TYPE: PPT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: branched-chain
 OTHER INFORMATION: heparin-binding peptide Lys Helix #1
 US-09-166-930A-5

Query Match 100.0%; Score 22; DB 4; Length 19;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 DB 15 KAAAK 19

RESULT 7
 PC-US92-06553-1
 Sequence 1; Application PC/US9206553
 GENERAL INFORMATION:
 APPLICANT: Katz M.D., David H
 APPLICANT: Matsushita M.D., SHO
 TITLE OF INVENTION: T-Cell Membrane Protein for B Cell IG
 NUMBER OF SEQUENCES: 2
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Grant L. Hubbard
 STREET: 336 North Seventh Street
 CITY: Phoenix
 STATE: AZ
 COUNTRY: USA
 ZIP: 85006
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: PC/US92/06553
 APPLICATION NUMBER: PC/US92/06553
 FILING DATE: 19920806
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/741,671
 FILING DATE: 07-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hubbard, Grant L
 REGISTRATION NUMBER: 24193

REFERENCE/DOCKET NUMBER: XE1032PC
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (602) 363-8782
 TELEFAX: (602) 365-5984
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 19 amino acids
 TYPE: AMINO ACID
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 ORIGINAL SOURCE: helper cell, hybridoma
 CELL LINE: MAT 1.15
 PUBLICATION INFORMATION:
 DOCUMENT NUMBER: US 07/41671
 FILING DATE: 07-AUG-1991
 PCT-US92-06553-1

Query Match 100.0%; Score 22; DB 5; Length 19;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KAAK 5
 Db 14 KAAK 18

RESULT 8
 US-08-473-344-57
 : Sequence 37, Application US/08473344
 : Patent No. 5763567
 : GENERAL INFORMATION:
 : APPLICANT: Little, Roger G
 : TITLE OF INVENTION: Biologically Active Peptides from
 : FUNCTIONAL DOMAINS OF BACTERICIDAL/PERMEABILITY-INCREASING
 : PROTEIN AND USES THEREOF
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Banner & Allegretti, Ltd.
 : STREET: 10 South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/473,344
 : FILING DATE: 7-JUN-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/306,473
 : FILING DATE: 15-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/209,762
 : FILING DATE: 11-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J.
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 93,1133-J
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 37:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 21 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : FEATURE KEY: misc feature
 : OTHER INFORMATION: BPI.47*

Query Match 100.0%; Score 22; DB 5; Length 19;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KAAK 5
 Db 14 KAAK 18

RESULT 8
 US-08-473-344-57
 : Sequence 37, Application US/08473344
 : Patent No. 5763567
 : GENERAL INFORMATION:
 : APPLICANT: Little, Roger G
 : TITLE OF INVENTION: Biologically Active Peptides from
 : FUNCTIONAL DOMAINS OF BACTERICIDAL/PERMEABILITY-INCREASING
 : PROTEIN AND USES THEREOF
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Banner & Allegretti, Ltd.
 : STREET: 10 South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/473,344
 : FILING DATE: 7-JUN-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/306,473
 : FILING DATE: 15-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/209,762
 : FILING DATE: 11-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J.
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 93,1133-J
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 37:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 21 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : FEATURE KEY: misc feature
 : OTHER INFORMATION: BPI.47*

TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE KEY: misc feature
 OTHER INFORMATION: BPI.45*
 US-08-473-344-57

Query Match 100.0%; Score 22; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KAAK 5
 Db 3 KAAK 7

RESULT 9
 US-08-473-344-58
 : Sequence 58, Application US/08473344
 : Patent No. 5763567
 : GENERAL INFORMATION:
 : APPLICANT: Little, Roger G
 : TITLE OF INVENTION: Biologically Active Peptides from
 : FUNCTIONAL DOMAINS OF BACTERICIDAL/PERMEABILITY-INCREASING
 : PROTEIN AND USES THEREOF
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Banner & Allegretti, Ltd.
 : STREET: 10 South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/473,344
 : FILING DATE: 7-JUN-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/306,473
 : FILING DATE: 15-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/209,762
 : FILING DATE: 11-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J.
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 93,1133-J
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 58:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 21 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : FEATURE KEY: misc feature
 : OTHER INFORMATION: BPI.47*

Query Match 100.0%; Score 22; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KAAK 5

RESULT 9
 US-08-473-344-58
 : Sequence 58, Application US/08473344
 : Patent No. 5763567
 : GENERAL INFORMATION:
 : APPLICANT: Little, Roger G
 : TITLE OF INVENTION: Biologically Active Peptides from
 : FUNCTIONAL DOMAINS OF BACTERICIDAL/PERMEABILITY-INCREASING
 : PROTEIN AND USES THEREOF
 : NUMBER OF SEQUENCES: 98
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Banner & Allegretti, Ltd.
 : STREET: 10 South Wacker Drive, Suite 3000
 : CITY: Chicago
 : STATE: Illinois
 : COUNTRY: USA
 : ZIP: 60606
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patent Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/473,344
 : FILING DATE: 7-JUN-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/306,473
 : FILING DATE: 15-SEP-1995
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/209,762
 : FILING DATE: 11-MAR-1995
 : ATTORNEY/AGENT INFORMATION:
 : NAME: McDonnell, John J.
 : REGISTRATION NUMBER: 26,949
 : REFERENCE/DOCKET NUMBER: 93,1133-J
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: 312-715-1000
 : TELEFAX: 312-715-1234
 : TELEX: 910-221-5317
 : INFORMATION FOR SEQ ID NO: 58:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 21 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: peptide
 : FEATURE KEY: misc feature
 : OTHER INFORMATION: BPI.47*

Query Match 100.0%; Score 22; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KAAK 5

```

Db 13 KAAK 17
|||||
RESULT 10
US-08-473-344-59
; Sequence 59, Application US/08473344
; Patent No. 5763567
; GENERAL INFORMATION:
; APPLICANT: Little, Roger G
; TITLE OF INVENTION: Biologically Active Peptides from
; TITLE OF INVENTION: Functional Domains of Bactericidal/Permeability-Increasing
; TITLE OF INVENTION: Protein and Uses Thereof
; NUMBER OF SEQUENCES: 99
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, Ltd.
; STREET: 10 South Wacker Drive, Suite 3009
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version 1.23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,344
; FILING DATE: 7-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/306,473
; FILING DATE: 15-SEP-1995
; PRIORITY INFORMATION:
; APPLICATION NUMBER: US 08/209,752
; FILING DATE: 11-MAR-1995
; NAME: McDonnell, John J.
; ATTORNEY/AGENT INFORMATION:
; REGISTRATION NUMBER: 26,949
; REFERENCE/DOCKET NUMBER: 93.1133-J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 21 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: "SPT.49"
US-08-473-344-59
Query Match 100.0%; Score 22; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 5
Db 3 KAAK 7
|||||
RESULT 11
PCT-US96-06224-1
; Sequence 1, Application PCT/US9606224
; GENERAL INFORMATION:
; APPLICANT: POMPLIANO, DAVID I
; APPLICANT: BRANHILL, DAVID
; APPLICANT: CUNNINGHAM, BARRY R
; APPLICANT: EL-SHERBEINI, MOHAMED
; APPLICANT: JONES, BRIAN
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 Camino de Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; APPLICANT: TRUSCA, DORINA
; TITLE OF INVENTION: A NOVEL BIOCHEMICAL INTERACTION IN
; TITLE OF INVENTION: BACTERIAL PATHOGENESIS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ELLIOTT KORSNEN
; STREET: 126 E. LINCOLN AVENUE, P.O. BOX 2000
; CITY: FAIRWAY
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/06224
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/437,253
; FILING DATE: 08-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: KORSNEN, ELLIOTT
; REGISTRATION NUMBER: 32,705
; REFERENCE/DOCKET NUMBER: 19455
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908 594-5493
; TELEFAX: 908 594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US96-06224-1
Query Match 100.0%; Score 22; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 5
Db 2 KAAK 6
|||||
RESULT 12
US-09-041-889-41
; Sequence 41, Application US/0901889
; Patent No. 6038661
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; APPLICANT: Cohavy, Offer
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 Camino de Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/837,058
APPLICATION NUMBER: US/08/041,889
CLASSIFICATION:
PRIORITY INFORMATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/837,058
FILING DATE: 11-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-PM 3006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 41:
SEQUENCE CHARACTERISTICS:
LENGTH: 55
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-041-889-41

Query Match 100.0%; Score 22; DB 3; Length 55;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 7 KAAK 11

RESULT 13
US-08-346-849-16
Sequence 16, Application US/08346849
Patent No. 5570483
GENERAL INFORMATION:
APPLICANT: Zhang, Shuguang
APPLICANT: Lockshin, Curtis
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,849
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28 DECEMBER 1992
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, DAVID E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-9540
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-293-284A-16

SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-346-849-16

Query Match 100.0%; Score 22; DB 1; Length 60;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 1 KAAK 5

RESULT 14
US-08-293-284A-16
Sequence 16, Application US/08293284A
Patent No. 5955343
GENERAL INFORMATION:
APPLICANT: Holmes, Todd
APPLICANT: Zhang, Shuguang
APPLICANT: Rich, Alexander
APPLICANT: Dipersio, C. Michael
TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY
TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES
TITLE OF INVENTION: THEREFOR
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/293,284A
FILING DATE: 22-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/973,326
FILING DATE: 28-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: BROOK, DAVID E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: MIT-6008A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-9540
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-293-284A-16

Query Match 100.0%; Score 22; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 65;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 1 KAAK 5

```

RESULT 15
US-09-041-889-38
; Sequence 38, Application US/09041889
; Patent No. 6033864
; GENERAL INFORMATION:
; APPLICANT: Braun, Jonathan
; TITLE OF INVENTION: Diagnosis, Prevention and Treatment of
; TITLE OF INVENTION: Ulcerative Colitis, and Clinical Subtypes Thereof, Using
; TITLE OF INVENTION: Microbial UC PANCA antigens
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIA TYPE: copy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/041.889
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/837,058
; FILING DATE: 12/28/1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-PM 3006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-041-889-38

```

```

Query Match 100.0%; Score 22; DB 3; Length 116;
Best local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 KAAK 5
DB 110 KAAK 114

```

```

Search completed: April 24, 2001, 16:36:23
Job time: 303 sec

```

GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OK protein - protein search, using sw model

Run on: April 24, 2001, 16:41:56 ; Search time 74.56 seconds
(without alignments)
4,609 Million cell updates/sec

Title: US-09-340-736-4
Perfect score: 22
Sequence: 1 KAAAK 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Method: 198801 seqs, 6872935 residues
Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PR_67: *
1: P11: *
2: P12: *
3: P13: *
4: P14: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	22	100.0	18	2	A61220	epsilon receptor m
2	22	100.0	20	2	S06149	photosystem I chl
3	22	100.0	38	2	S21173	photosystem I chl
4	22	100.0	45	2	A05163	antifreeze protein
5	22	100.0	58	2	A71527	probable S21 ribos
6	22	100.0	58	2	F72128	ribosomal protein
7	22	100.0	58	2	F81683	ribosomal protein
8	22	100.0	82	2	T30977	hypothetical prote
9	22	100.0	82	1	S27135	ribosomal protein
10	22	100.0	87	1	S27135	ribosomal protein
11	22	100.0	76	2	T45825	ribosomal protein
12	22	100.0	77	2	S45375	ribosomal protein
13	22	100.0	81	1	JN0452	adenoregulin prsu
14	22	100.0	82	2	G75426	hypothetical prote
15	22	100.0	87	2	T27141	hypothetical prote
16	22	100.0	93	2	F70978	hypothetical prote
17	22	100.0	100	2	A33310	nonhistone chromos
18	22	100.0	101	2	S65494	sperm-specific pro
19	22	100.0	105	1	JN0483	ribosomal protein
20	22	100.0	105	2	T09720	ribosomal protein
21	22	100.0	110	2	T41725	ribosomal protein
22	22	100.0	111	2	S41745	photosystem I chl
23	22	100.0	115	1	F13594	histone H1-like pr
24	22	100.0	125	2	A39396	probable histone-1
25	22	100.0	125	2	A71477	histone H1-like pr
26	22	100.0	125	2	B81739	hypothetical prote
27	22	100.0	132	2	F82800	hypothetical prote
28	22	100.0	136	2	J00757	histone H3 - stagh
29	22	100.0	142	2	C49652	hypothetical prote

high mobility grou
histone H2A.2 - wh
RNA methyltransfe
RNA methyltransfe
hypothetical prote
Lsu ribosomal prot
probable ribosomal
histone H1 - Tetra
H+-transporting AT
myosin light chain
transcription regu
histone H1 homolog
histone H1-delta
hypothetical prote
histone H1 - trout
hypothetical prote

ALIGNMENTS

RESULT 1
A61220
epsilon-receptor modulating protein (BC 3.4.21.1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 21-Mar-1995
C:Accession: A61220
R:Matsumita, S.; Katz, D.H.
Cell. Immunol. 137, 252-259, 1991
A:Title: The murine epsilon receptor modulating protein: a novel serine protease whic
A:Reference number: A61220; MUID:91356570
A:Accession: A61220
A:Residues type: protein
A:Residues: 1-18 CAAK
C:Comment: This serine proteinase from a T cell hybridoma does not reduce levels of C
C:Keywords: hydrolase; serine proteinase

Query Match 100.0%; Score 22; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAAK 5

Db 13 KAAAK 17

RESULT 2
S06149
photosystem I chain III - barley (fragment)
C:Species: Hordeum vulgare (barley)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 18-Jun-1993
C:Accession: S06149
R:Arundon, S.; Vanstein, A.; Thornber, J.P.
Biolett. 256, 135-139, 1989
A:Title: Correlation of published amino acid sequences for photosystem I polypep
A:Reference number: S06149; MUID:30033290
A:Accession: S06149
A:Molecule type: protein
A:Residues: 1-20 KAAK
A:Note: the authors identified this protein as light-harvesting complex I 17K protein
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 100.0%; Score 22; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAAK 5

Db 12 KAAAK 16


```

RESULT 3
Phocystern I chain IV precursor - spinach
C:Accession: S23173
C>Date: 19-Sep-1997 #sequence_revision 19-Mar-1997 #text_change 07-May-1999.
C:Accession: S23173
R:Laigrette, B.; Vallon, O.
Eur. J. Biochem. 205, 1175-1185, 1992
A:Title: Purification and membrane topology of PSI-D and PSI-E, two subunits of the phot
A:Reference number: S23173; MUID:92249324
A:Accession: S23173
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-38 <IAG>

Query Match 100.0%; Score 22; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
Db 20 KAAK 24

RESULT 4
antifreeze protein SS-8 - shorthorn sculpin
C:Species: Myoxocephalus scorpius (shorthorn sculpin, daddy sculpin)
C>Date: 05-Jun-1997 #sequence_revision 05-Jun-1997 #text_change 24-Nov-1999
C:Accession: A05163
R:Hew, C.L.; Joshi, S.; Wang, N.C.; Rao, M.H.; Ananthanarayanan, V.S.
Eur. J. Biochem. 151, 167-172, 1985
A:Title: Structures of shorthorn sculpin antifreeze polypeptides.
A:Reference number: A05163; MUID:85285003
A:Accession: A05163
A:Molecule type: protein
A:Residues: 1-45 <HEW>
C:Superfamily: antifreeze protein
C:Keywords: antifreeze; blocked amino end; plasma
F:9-45/Region: alanine-rich
F:1/Modified site: blocked amino end (Met) #status experimental

Query Match 100.0%; Score 22; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
Db 30 KAAK 34

RESULT 5
probable S21 ribosomal protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)
C:Species: Chlamydia trachomatis
C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 08-Oct-1999
C:Accession: A71527
R:Stephens, R.S.; Alaman, S.; Lamell, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,
A:Title: 28S rRNA of Chlamydia trachomatis serotype D.
A:Reference number: A71527; MUID:9000809
A:Accession: A71527
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-58 <ARN>
A:Cross-references: GB:AE001307; GB:AE001273; NID:g3328757; PIDN:AAC67937.1; PID:g332876
A:Experimental source: serotype D, strain UW-3/Cx
C:Genetics: rs21
A:Gene: rs21

```

```

Query Match 100.0%; Score 22; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
Db 50 KAAK 54

RESULT 6
ribosomal protein S21 CP0745 [imported] - Chlamydia pneumoniae (strains CWL029 an
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
R:Accession: F72128; MUID:831342
R:DeBooy, R.; Nelson, W.; DeBooy, R.; Kolonay, J.; McClarty, G.; Salzbe
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39
A:Reference number: AB1500; MUID:20150255
A:Accession: G81542
A:Molecule type: DNA
A:Residues: 1-58 <REA>
A:Cross-references: GB:AE002233; GB:AE002161; NID:g7189657; PIDN:AAF38550.1; PID:g718
A:Experimental source: strain AR39, HL cells
C:Genetics: rs21; CP0745

Query Match 100.0%; Score 22; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
Db 50 KAAK 54

RESULT 7
ribosomal protein S21 TC0620 [imported] - Chlamydia muridarum (strain Nigg5)
C:Species: Chlamydia muridarum, Chlamydia trachomatis MOPn
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
R:Accession: B81683
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heideberg, J.F.; White, O.; Hicke
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOPn and Chlamydia pneumoniae AR39
A:Reference number: AB1500; MUID:20150255
A:Accession: B81683
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-58 <TC0620>
A:Cross-references: GB:AE002330; GB:AE002160; NID:g7190649; PIDN:AAF39451.1; PID:g719
A:Experimental source: strain Nigg (MOPn)
C:Genetics:
A:Gene: TC0620

Query Match 100.0%; Score 22; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5

```

```
Db 50 KAAK 54
|||||
RESULT 8
S16270
A:Accession: S16270
A:Title: Ribosomal protein L35 - Synecocystis sp. (strain POC 6803)
C:Species: Synecocystis sp.
A:Variety: POC 6803
C:Date: 15-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Aemizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Reference number: S76270
A:Accession: S76270
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys
A:Reference number: S74322; MID: 97061201
A:Accession: S76270
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-67 <KAN>
A:Cross-references: ENBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA010122.1; PID:g100
A:Note: the nucleotide sequence was submitted to the ENBL Data Library, June 1996
C:Genetics:
A:Gene: rpl35
A:Superfamily: Escherichia coli ribosomal protein L35
A:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 22; DB 2; Length 67;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
|||||
Db 46 KAAK 50

RESULT 9
P5XT35
A:Accession: P5XT35
A:Title: Ribosomal protein L35 - Cyanophora paradoxa cyanelle
C:Species: Cyanelle Cyanophora paradoxa
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 22-Jun-1999
R:Bryant, D.A.; Stirewalt, V.L.
FEBS Lett. 259, 273-280, 1990
A:Reference number: S07070; T06849
A:Accession: S07070; T06849
A:Title: The cyanelle genome of Cyanophora paradoxa encodes ribosomal proteins not encod
A:Reference number: S07067; MID: 30073561
A:Accession: S07070
A:Molecule type: DNA
A:Residues: 1-65 <BY>
A:Cross-references: ENBL:X17063; NID:g11291; PIDN:CA34907.1; PID:g11292
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohmert, H.J.; Bryant, D.A.
submitted to the ENBL Data Library, July 1995
A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.
A:Reference number: S06849
A:Accession: S06849
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-65 <STI>
A:Cross-references: ENBL:U30821; NID:g1016083; PIDN:AAA81152.1; PID:g1016105
A:Experimental source: strain Pringsheim LB555
C:Genetics:
A:Gene: rpl35; rpm1
A:Map position: 27
A:Genome: Cyanelle
A:Superfamily: Escherichia coli ribosomal protein L35
A:Keywords: cyanelle; protein biosynthesis; ribosome

Query Match 100.0%; Score 22; DB 1; Length 65;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
|||||

Db 8 KAAK 12
|||||
RESULT 10
S16270
A:Accession: S16270
A:Title: Ribosomal protein L35 - Synecocystis sp. (strain POC 6803)
C:Species: Synecocystis sp.
A:Variety: POC 6803
C:Date: 15-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Aemizu, E.; Nakamura, Y.; Miyajima,
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
DNA Res. 3, 109-136, 1996
A:Reference number: S76270
A:Accession: S76270
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys
A:Reference number: S74322; MID: 97061201
A:Accession: S76270
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-67 <KAN>
A:Cross-references: ENBL:D64000; GB:AB001339; NID:g1001484; PIDN:BA010122.1; PID:g100
A:Note: the nucleotide sequence was submitted to the ENBL Data Library, June 1996
C:Genetics:
A:Gene: rpl35
A:Superfamily: Escherichia coli ribosomal protein L35
A:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 22; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
|||||
Db 12 KAAK 16

RESULT 11
145885
A:Accession: 145885
A:Title: Bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1995 #sequence_revision 15-Oct-1996 #text_change 13-Aug-1999
R:Stirewalt, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohmert, H.J.; Bryant, D.A.
submitted to the ENBL Data Library, July 1995
A:Reference number: I45885; MID: 85059254
A:Accession: I45885
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-76 <ROS>
A:Cross-references: GB:X31891; NID:g153008; PIDN:AAA96416.1; PID:g552319
A:Experimental source: strain Pringsheim LB555
A:Introns: 20/1, 58/1
C:Superfamily: elastin

Query Match 100.0%; Score 22; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
|||||
Db 12 KAAK 16

RESULT 12
S45375
A:Accession: S45375
A:Title: Ribosomal protein S4 - common tobacco (fragment)
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 06-Jan-1995 #sequence_revision 07-Feb-1997 #text_change 13-Aug-1999
C:Accession: S45375
```

R.Marty, I.; Brugidou, C.; Chartier, Y.; Meyer, Y.
 Plant J. 4, 265-278, 1993
 A>Title: Growth-related gene expression in Nicotiana tabacum mesophyll protoplasts.
 A:Reference number: 845375
 A:Accession: 845375
 A:Molecule type: mRNA
 A:Residues: 1-77 <MT>
 A:Cross-references: EMBL:Z14085; NID:9443959; PIDN:CAA76463.1; PID:9443960
 A:Superfamily: rat ribosomal protein S9
 C:Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 22; DB 2; Length 77;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 |||||
 DB 60 KAAAK 64

RESULT 13

Adenoregulin precursor - two-colored leaf frog
 A:Alternate names: dermaseptin b II
 C:Species: Phyllomedusa bicolor (two-colored leaf frog)
 C:Accession: J00462; EMBL:Z14171; NID:9395932; PIDN:CAA49763.1; PID:9395932
 R:Baiche, M.; Duacanal, F.; Rajenassae, E.; Boullain, J.C.; Venez, A.; Nicolas, P.
 Biochem. Biophys. Res. Commun. 191, 983-990, 1993
 A>Title: Molecular cloning of a cDNA encoding the precursor of adenoregulin from frog sk
 A:Reference number: JN0462; NID:93221546
 A:Accession: JN0462
 A:Molecule type: mRNA
 A:Residues: 1-81 <MT>
 A:Cross-references: EMBL:X70278; NID:9395931; PIDN:CAA49763.1; PID:9395932
 A:Experimental source: skin
 A>Note: the authors translated the codon AAA for residue 26 as Ieu
 A:Accession: J00462; EMBL:Z14171; NID:9395932; PIDN:CAA49763.1; PID:9395932
 J:8101 Chn. 269 17845-17852, 1993
 A>Title: Precursors of vertebrate peptide antibiotics dermaseptin b and adenoregulin hav

A:Reference number: A54897; NID:94299491

A:Accession: A54897
 A:Molecule type: mRNA
 A:Residues: 1-81 <MT>
 A:Cross-references: GB:X70278; NID:9395931; PIDN:CAA49763.1; PID:9395932
 R:Daly, J.W.; Caceres, J.; Monti, R.W.; Gusovsky, F.; Mocs Jr., M.; Seamon, K.B.; Milton,
 Proc. Natl. Acad. Sci. U.S.A. 89, 10960-10963, 1992
 A>Title: Frog secretions and hunting magic in the upper Amazon: identification of a pep
 A:Accession: M4171; NID:93060363

A:Molecule type: protein
 A:Residues: 46-78 <MT>
 C:Superfamily: dermaseptin precursor; dermorphin precursor amino-terminal homology
 C:Keywords: antibiotic; antifungal; skin.
 F:1-45/Domain: dermorphin precursor amino-terminal homology <DER>
 F:1-22/Domain: signal sequence #status predicted <SIG>
 F:23-45/Domain: propeptide #status predicted <PRO>
 F:46-78/Product: adenoregulin #status experimental <MT>

Query Match 100.0%; Score 22; DB 1; Length 81;
 Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 |||||
 DB 60 KAAAK 64

RESULT 14

G75426

Hypothetical protein - Deinococcus radiodurans (strain RL)
 C:Species: Deinococcus radiodurans
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: G75426
 R:Witte, C.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 S. M. Shen, W.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, L.; Kaleski, C.;
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL.
 A:Reference number: A75250; NID:20036896
 A:Accession: G75426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-82 <MT>
 A:Cross-references: GB:AE001967; GB:AE000513; NID:96458915; PIDN:AAF10756.1; PID:9645

A:Experimental source: strain RL
 C:Accession: G75426
 R:Witte, C.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 S. M. Shen, W.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, L.; Kaleski, C.;
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans RL.
 A:Reference number: A75250; NID:20036896
 A:Accession: G75426
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-82 <MT>
 A:Cross-references: GB:AE001967; GB:AE000513; NID:96458915; PIDN:AAF10756.1; PID:9645

Query Match 100.0%; Score 22; DB 2; Length 82;

Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 |||||
 DB 62 KAAAK 66

RESULT 15

Hypothetical protein Y53F4A.b - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27141
 R:Smyle, R.
 submitted to the EMBL Data Library, September 1999

A:Reference number: T27141
 A:Accession: T27141
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-87 <MT>
 A:Cross-references: EMBL:AL110494; NID:el542292; PIDN:CA554462.1; CESP:Y53F4A.b
 A:Experimental source: clone Y53F4A
 C:Genetics:
 A:Gene: CESP:Y53F4A.b
 A:Introns: 76/3

Query Match 100.0%; Score 22; DB 2; Length 87;

Best Local Similarity 100.0%; Pred. No. 1.le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAAK 5
 |||||
 DB 26 KAAAK 30

Search completed: April 24, 2001, 16:41:58

Job time: 463 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:42:47 ; Search time 44.88 seconds
(without alignments)
3.816 Million cell updates/sec

Title: US-09-340-736-4

Perfect score: 22

Sequence: 1 KRAAK 5

Scoring table: BLOSUM62

Gap: 10.0 , Gapext 0.5

Method: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the best hit being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Match Length DB ID Description

1 22 100.0 27 1 DMS4_PHYSA P0280 Phylomedusa

2 22 100.0 45 1 KRAAK 5 P0350 Xenopus lae

3 22 100.0 45 1 KRAAK 5 P0350 Xenopus lae

4 22 100.0 58 1 RS21_CHLPH P0350 Xenopus lae

5 22 100.0 58 1 RS21_CHLPH P0350 Xenopus lae

6 22 100.0 65 1 K335_CVAPA P0350 Xenopus lae

7 22 100.0 67 1 RL35_SVW3 P0350 Xenopus lae

8 22 100.0 74 1 RL35_SVW3 P0350 Xenopus lae

9 22 100.0 77 1 RS2_TORAC P0350 Xenopus lae

10 22 100.0 80 1 RL35_SVW3 P0350 Xenopus lae

11 22 100.0 81 1 DMS2_PHEI P0350 Xenopus lae

12 22 100.0 81 1 DMS2_PHEI P0350 Xenopus lae

13 22 100.0 99 1 K335_CVAPA P0350 Xenopus lae

14 22 100.0 99 1 K335_CVAPA P0350 Xenopus lae

15 22 100.0 104 1 RL36_RAT P0350 Xenopus lae

16 22 100.0 110 1 RS25_DICDI P0350 Xenopus lae

17 22 100.0 111 1 RL36_RAT P0350 Xenopus lae

18 22 100.0 125 1 RL36_RAT P0350 Xenopus lae

19 22 100.0 125 1 RL36_RAT P0350 Xenopus lae

20 22 100.0 135 1 H3_ACRFO P0350 Xenopus lae

21 22 100.0 154 1 KRAAK 5 P0350 Xenopus lae

22 22 100.0 161 1 RL36_RAT P0350 Xenopus lae

23 22 100.0 161 1 RL36_RAT P0350 Xenopus lae

24 22 100.0 164 1 RL11_PVHRO P0350 Xenopus lae

25 22 100.0 164 1 RL11_PVHRO P0350 Xenopus lae

26 22 100.0 165 1 RL11_PVHRO P0350 Xenopus lae

27 22 100.0 165 1 RL11_PVHRO P0350 Xenopus lae

28 22 100.0 172 1 ATPE_BACME P0350 Xenopus lae

29 22 100.0 185 1 RL36_RAT P0350 Xenopus lae

30 22 100.0 189 1 RL36_RAT P0350 Xenopus lae

31 22 100.0 194 1 RL36_RAT P0350 Xenopus lae

32 22 100.0 194 1 RL36_RAT P0350 Xenopus lae

33 22 100.0 199 1 RL36_RAT P0350 Xenopus lae

P13375 strongyloce
P06350 oncorhynch
P06892 xenopus lae
P15869 strongyloce
P04730 chylomydia
P04730 chylomydia
P15866 xenopus lae
P07796 strongyloce
P09987 gallus gall
P08284 gallus gall
P08288 gallus gall
P10412 homo sapien

34 22 100.0 205 1 H1E_STRPU
35 22 100.0 206 1 H1E_STRPU
36 22 100.0 209 1 H1E_STRPU
37 22 100.0 211 1 H1E_STRPU
38 22 100.0 213 1 H1E_STRPU
39 22 100.0 213 1 H1E_STRPU
40 22 100.0 216 1 H1E_STRPU
41 22 100.0 217 1 H1E_STRPU
42 22 100.0 217 1 H1E_STRPU
43 22 100.0 218 1 H1E_STRPU
44 22 100.0 218 1 H1E_STRPU
45 22 100.0 218 1 H1E_STRPU

ALIGNMENTS

RESULT 1

DMS4_PHYSA STANDARD: PRT: 27 AA.

AC P0280; 1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DR 01-FEB-1994 (Rel. 28, Last annotation update)

DE DMS4_PHYSA (DS IV)

OS Phylomedusa sauvagei (Sauvage's leaf frog)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;

CC Phylomedusa

OK NCBI_TaxID=8395;

RR (1)

REQUENCE

RS TSSUP-SKIN

FX MEDLINE=94135686; PubMed=8306981;

FX Mor A., Nicolas P.

RT "Isolation and structure of novel defensive peptides from frog skin."

RL Eur. J. Biochem. 219:145-154(1994).

CC -1- FUNCTION: POSSESSES A POTENT ANTIMICROBIAL ACTIVITY AGAINST

CC BACTERIA FUNGI AND PROTOZOA. PROBABLY ACTS BY DISTURBING MEMBRANE

CC -1- FUNCTIONS WITH ITS AMPHIPHILIC STRUCTURE.

CC -1- TISSUE/LOC: SECRETED.

KW Antibiotic; Peptide; Multipeptide family; Amphibian skin.

SQ SEQUENCE 27 AA; 2779 WG; 43C94D2DC19721A8 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 27;

Best Local Similarity 100.0%; Frad. No. 17;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KRAAK 5

DB 12 KRAAK 16

RESULT 2

DMS5_PHYSA STANDARD: PRT: 29 AA.

AC P0281; 1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DR 01-FEB-1994 (Rel. 28, Last annotation update)

DE DMS5_PHYSA (DS V)

OS Phylomedusa sauvagei (Sauvage's leaf frog)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

CC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;

CC Phylomedusa

OK NCBI_TaxID=8395;

RR (1)

RESEQUENCE

RS TSSUP-SKIN

FX MEDLINE=94135686; PubMed=8306981;

FX Mor A., Nicolas P.

RT "Isolation and structure of novel defensive peptides from frog skin.";
 RL Eur. J. Biochem. 219:145-154 (1994).
 CC -1- FUNCTION: POSSESSES A POTENT ANTIMICROBIAL ACTIVITY AGAINST
 CC FUNGICIDAL AND PROZOAL PROPERTIES. ACTS BY DISTURBING MEMBRANE
 CC FUNCTION. DISEASE: CHLAMYDIA. STRUCTURE:
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- TISSUE SPECIFICITY: SKIN.
 KW Antibiotic; Fungicide; Multigene family; Amphibian skin.
 SQ SEQUENCE 29 AA; 2840 MW; 540A4971FC5B506 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 29;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 15 KAAK 19

RESULT 3
 ID RS21_MYOSC STANDARD: PRT; 45 AA.
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 01-MAR-1987 (Rel. 04, Last sequence update)
 DE ANTI-FREEZE PEPTIDE (S-G)
 OS Myoxocephalus scorpius (Shorthorn sculpin) (Daddy sculpin).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Neuteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neuteleostei;
 CC Acanthomorpha; Acanthopterygii; Percomorphi; Scorpaeniformes;
 CC Cottolidae; Cottidae; Myoxocephalus.
 OX NCBI_Taxid=8097;
 RN [1]

RP SEQUENCE 45 AA; 4096 MW; 260C0BC563B6878 CRC64;
 RA "Structures of shorthorn sculpin antifreeze polypeptides.";
 RT Eur. J. Biochem. 151:167-172 (1985).
 RL -1- FUNCTION: ANTIFREEZE PROTEINS LOWER THE BLOOD FREEZING POINT.
 CC -1- SIMILARITY: BELONGS TO THE TYPE-I AFP FAMILY. TYPE 1 AFP ARE
 CC ALANINE-RICH, AMPHIPHILIC AND ALPHA-HELICAL.
 DR HSSP; P04002; IWFA.
 DR Antifreeze protein; Repeat.
 FT MOLES 1 1 BLOCKED.
 FT REPEAT 22 21
 FT REPEAT 34 45
 FT REPEAT 34 45

Query Match 100.0%; Score 22; DB 1; Length 45;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 30 KAAK 34

RESULT 4
 ID RS21_CHLPN STANDARD: PRT; 58 AA.
 AC Q949F0; Q9J0K3;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE 30S RIBOSOMAL PROTEIN S21.
 GN RPSU OR RS21 OR CPN031 OR CP0745.
 OS Chlamydia trachomatis (Chlamydia pneumoniae).
 CC Bacteria; Chlamydiales; Chlamydiales; Chlamydia.

OX NCBI_Taxid=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92106066; PubMed=10122386;
 RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RA "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RA Nat. Genet. 21:385-389 (1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STRAIN=AR39;
 RA MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Otterback T., Barry K., Bass S.,
 RA Linder K., Weidman J., Khouri H., Craven B., Roman C., Dodson R.,
 RA Evans T., Nelson M., DeBoy K., Kolony J., McClarty G., Salzberg S.L.,
 RA "Genome sequences of Chlamydia trachomatis MOpn and Chlamydia
 RA pneumoniae AR39.";
 RA Nucleic Acids Res. 28:1397-1405 (2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA STRAIN=J138;
 RA MEDLINE=20303049; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kubota S., Nakazawa T.,
 RA "Complete genome sequence of Chlamydia pneumoniae J138
 RA from Japan and ChLO29 from USA.";
 RA Nucleic Acids Res. 28:2311-2314 (2000).
 CC -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC
 CC EXBL; AB01588; AAD18184.1;
 DR EXBL; AB02233; AAF38550.1;
 DR EXBL; AF002545; BAA98243.1;
 DR TIGR; CF0745;
 DR InterPro: IPR001911;
 DR Pfam: PF01165; Ribosomal_S21; 1.
 DR PRINIS: PR00976; RIBOSOMAL_S21;
 DR PROSITE: PS01161; RIBOSOMAL_S21; FALSE_NEG.
 RN RIBOSOMAL PROTEIN.
 KW SEQUENCE 58 AA; 6636 MW; 74118A525600177 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 58;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 50 KAAK 54

RESULT 5
 ID RS21_CHLTR STANDARD: PRT; 58 AA.
 AC Q84346;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE 30S RIBOSOMAL PROTEIN S21.
 GN RPSU OR RS21 OR CT342 OR TC0620.
 OS Chlamydia trachomatis, and
 OS Chlamydia muridarum.
 CC Bacteria; Chlamydiales; Chlamydiales; Chlamydia.

KW Ribosomal protein.
SQ SEQUENCE 67 AA; 7891 MW; E43B50E79B5F294 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 67;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 8 KAAK 12

RESULT 8
H16L_TRYCR STANDARD; PRT: 74 AA.
AC P40270; P40272;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HISTONE H16L/H1.M.
RN Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
NCBI_TaxID=5693;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-TULAHUEN 2;
RX MEDLINE=95059220; PubMed=7969272;
RA Aslund L., Carlsson L., Henriksson J., Rydaaker M., Toro G.C.,
RA Galanti N., Petersson U.,
RT A gene family encoding heterogeneous histone H1 proteins in
RI Trypanosoma cruzi.
RI Mol. Biochem. Parasitol. 65:317-330(1994).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; L27116; AA066480.1;
DR EMBL; L27118; AA066480.1;
DR EMBL; L27119; AA066483.1;
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
SQ SEQUENCE 74 AA; 7688 MW; 036A76C72BE0249B CRC64;

Query Match 100.0%; Score 22; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 15 KAAK 19

RESULT 9
RS9_TOBAC STANDARD; PRT: 77 AA.
AC P49214;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE 40S RIBOSOMAL PROTEIN S9 (S4) (FRAGMENT).
RN Nicotiana tabacum (Common tobacco).
RN Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;
OC Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN (1)
RP SEQUENCE FROM N.A.

RC STRAIN-CV 19;
RX MEDLINE=94035181; PubMed=8230482;
RA Verity L., Rudigon C., Chartier Y., Meyer Y.,
RT "Growth-related gene expression in Nicotiana tabacum mesophyll
protoplasts";
RL Plant J. 4:265-278(1993).
CC -1- SIMILARITY: BELONGS TO THE SAP FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; Z14085; CA078463.1;
DR InterPro: IPR001912;
DR InterPro: IPR002942;
DR Pfam: PF01479; S4; 1;
DR PROSITE; PS00632; RIBOSOMAL_S4; PARTIAL.
KW Ribosomal protein; Nuclear protein; DNA-binding; Multigene family.
FT NON_TER 1
SQ SEQUENCE 77 AA; 8557 MW; 072E8E031A7D2582 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 77;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 50 KAAK 64

RESULT 10
H16L_TRYCR STANDARD; PRT: 80 AA.
AC P40273;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HISTONE H16L/H1.M.
RN Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
NCBI_TaxID=5693;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-TULAHUEN 2;
RX MEDLINE=95059220; PubMed=7969272;
RA Aslund L., Carlsson L., Henriksson J., Rydaaker M., Toro G.C.,
RA Galanti N., Petersson U.,
RT A gene family encoding heterogeneous histone H1 proteins in
RI Trypanosoma cruzi.
RI Mol. Biochem. Parasitol. 65:317-330(1994).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR EMBL; L27116; AA066480.1;
DR EMBL; L27118; AA066480.1;
DR EMBL; L27119; AA066483.1;
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
SQ SEQUENCE 80 AA; 8271 MW; 8E83DA30F0A5D31 CRC64;

Query Match 100.0%; Score 22; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
DB 15 KAAK 19
RESULT 11
ID DMS2_PHYBI STANDARD: PRT: 81 AA.
AC P31107; P80283;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last annotation update)
OS DMS2_PHYBI (DERMASEPTIN B1) (cos.)
OC Amphibia: Batrachia: Chordata: Vertebrata: Euteleostomi:
OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:
OC Phylomedusa.
CC [1]
NCBI_TaxID=8393;
PP SEQUENCE FROM N.A.
RX TISSUE=skin;
RX MEDLINE=9321546; PubMed=846657;
RA Aniche M., Ducancel F., Lajenne E., Boulain J.-C., Menez A.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [2]
PP SEQUENCE OF 46-78.
RX TISSUE=skin;
RX MEDLINE=9306663; PubMed=1438301;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [3]
PP SYNTHESIS, AND CHARACTERIZATION.
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [4]
PP FUNCTION: ENHANCES BINDING OF AGONISTS TO A1 ADENOSINE RECEPTORS.
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [5]
PP FUNCTION: POSSIBLES A POTENT ANTIMICROBIAL ACTIVITY AGAINST
BACTERIA, FUNGI AND PROTOZOA. PROBABLY ACTS BY DISTURBING MEMBRANE
FUNCTIONS WITH ITS AMPHIPATHIC STRUCTURE.
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [6]
PP SUBCELLULAR LOCATION: SECRETED.
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [7]
PP DISEASE: AFFECTS HUMAN BEHAVIOR ELICITING PROFOUND MALADISE,
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [8]
PP SIMILARITY: TO P. SUVADEI DERMASEPTIN.
RX TISSUE=skin;
RX MEDLINE=9413986; PubMed=8306981;
RA Dely J.W., Caecere V., Xomi R.W., Guevsky F., Moos M. Jr.,
RX "Molecular cloning of a cDNA encoding the precursor of adenosin
RT from frog skin. Relationships with the vertebrate defensive peptides,
RT dermaseptins.";
RL Biochem. Biophys. Res. Commun. 191:983-990(1993).
RN [9]
PP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC [10]
DR EMBL: X70278; CAA49763.1;
DR PIR: A44171; A44171.
DR UNIPROT: P44171; A44171.
DR Antibiotic; Fungicide; Multigene family: Amphibian skin: Signal;
KW Cleavage on pair of basic residues.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 45 POTENTIAL.
FT CHAIN 46 78 ADENOREGULIN.
FT PROPEP 79 81
FT PROPEP 81 81
PP SEQUENCE 81 AA: 8444 MW: C36AD8429418272D CRC64;
QY 1 KAAK 5
DB 60 KAAK 64
RESULT 12
ID H162_PHYBI STANDARD: PRT: 90 AA.
AC P40274;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HISTONE H1.M6.2.
OS Trypanosoma cruzi.
CC Trypanosoma cruzi.
CC Eukaryota: Euglenozoa: Kinetoplastida: Trypanosomatidae: Trypanosoma.
CC NCBI_TaxID=5693;
RX TISSUE=skin;
RX MEDLINE=9505220; PubMed=7569272;
RA Aslund L., Carlsson L., Henriksson J., Rydmark M., Toro G.C.,
RA Galanti M., Pettersson U.,
RX "A gene family encoding heterogeneous histone H1 proteins in
RT Trypanosoma cruzi.";
RL Mol. Biochem. Parasitol. 65:317-330(1994).
RN [1]
PP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC [2]
DR EMBL: L27119; AAA66484.1;
DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
PP SEQUENCE 90 AA: 9236 MW: 89950A95D5960C CRC64;
QY 1 KAAK 5
DB 15 KAAK 19
RESULT 13
ID H162_PHYBI STANDARD: PRT: 99 AA.
AC P40274;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HISTONE H1.M6.2.
OS Trypanosoma cruzi.
CC Trypanosoma cruzi.
CC Eukaryota: Euglenozoa: Kinetoplastida: Trypanosomatidae: Trypanosoma.
CC NCBI_TaxID=5693;
RX TISSUE=skin;
RX MEDLINE=9505220; PubMed=7569272;
RA Aslund L., Carlsson L., Henriksson J., Rydmark M., Toro G.C.,
RA Galanti M., Pettersson U.,
RX "A gene family encoding heterogeneous histone H1 proteins in
RT Trypanosoma cruzi.";
RL Mol. Biochem. Parasitol. 65:317-330(1994).
RN [1]
PP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC [2]
DR EMBL: L27119; AAA66484.1;
DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
PP SEQUENCE 90 AA: 9236 MW: 89950A95D5960C CRC64;
QY 1 KAAK 5
DB 15 KAAK 19
RESULT 14
ID H162_PHYBI STANDARD: PRT: 99 AA.
AC P40274;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HISTONE H1.M6.2.
OS Trypanosoma cruzi.
CC Trypanosoma cruzi.
CC Eukaryota: Euglenozoa: Kinetoplastida: Trypanosomatidae: Trypanosoma.
CC NCBI_TaxID=5693;
RX TISSUE=skin;
RX MEDLINE=9505220; PubMed=7569272;
RA Aslund L., Carlsson L., Henriksson J., Rydmark M., Toro G.C.,
RA Galanti M., Pettersson U.,
RX "A gene family encoding heterogeneous histone H1 proteins in
RT Trypanosoma cruzi.";
RL Mol. Biochem. Parasitol. 65:317-330(1994).
RN [1]
PP THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC [2]
DR EMBL: L27119; AAA66484.1;
DR Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
KW Chromosomal protein; Nuclear protein; DNA-binding; Multigene family.
PP SEQUENCE 90 AA: 9236 MW: 89950A95D5960C CRC64;

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EBLI_X06264; CNA8345.1; -
DR FAR_X06465; CNA86;
DR TAPAP0000509; -
DR Pfam: PF01158; Ribosomal_L36e; 1.
DR PROSITE: PS01190; RIBOSOMAL_L36e; 1.
KW Ribosomal protein.
FT INIT_MET 0 0
SQ SEQUENCE 104 AA; 12136 MW; A27B76C348FE9E1A CRC64;

Query Match 100.0%; Score 22; DB 1; Length 104;
Identical Similarity 100.0%; Positives 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 KAAK S
|||||
98 KAAAY 102

Search completed: April 24, 2001, 16:42:48
Job time: 454 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:40:30 ; Search time 125.5 Seconds
(without alignments)
4,670 Million cell updates/sec

Title: US-09-340-736-4

Perfect score: 22

Sequence: 1 KAAAK 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Ched: 374700 segs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPTEMBL15.*

2: sp_archaea.*

3: sp_bacteria.*

4: sp_fungi.*

5: sp_human.*

6: sp_mammal.*

7: sp_mbc.*

8: sp_morganelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_protist.*

12: sp_unclassified.*

13: sp_vertebrate.*

14: sp_virus.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	Description
1	22	100.0	13 10 Q95016	Q95016 oryza sativ
2	22	100.0	24 10 Q95825	Q95825 Spinacia ol
3	22	100.0	58 2 Q95825	Q95825 Spinacia ol
4	22	100.0	58 2 Q95825	Q95825 Spinacia ol
5	22	100.0	60 5 Q93393	Q93393 casenohabdi
6	22	100.0	61 5 Q9NFJ7	Q9NFJ7 trypanosoma
7	22	100.0	61 5 Q9NFJ5	Q9NFJ5 trypanosoma
8	22	100.0	62 5 Q17536	Q17536 casenohabdi
9	22	100.0	66 5 Q9NFJ3	Q9NFJ3 trypanosoma
10	22	100.0	71 5 Q9NFJ8	Q9NFJ8 trypanosoma
11	22	100.0	71 5 Q9NFJ5	Q9NFJ5 trypanosoma
12	22	100.0	71 5 Q9NFJ3	Q9NFJ3 trypanosoma
13	22	100.0	72 2 P96100	P96100 thioabacillu
14	22	100.0	72 2 Q9NFJ0	Q9NFJ0 trypanosoma
15	22	100.0	72 2 Q9NFJ0	Q9NFJ0 trypanosoma
16	22	100.0	75 5 Q9NFJ6	Q9NFJ6 trypanosoma
17	22	100.0	75 5 Q9NFJ4	Q9NFJ4 trypanosoma
18	22	100.0	76 5 Q9NFJ2	Q9NFJ2 trypanosoma
19	22	100.0	76 5 Q9NFJ1	Q9NFJ1 trypanosoma

ALIGNMENTS

RESULT 1

Q95016 PRELIMINARY; PRT: 15 AA.

DC Q95016

DT 01-MAY-2000 (TRENDEL, 13, Created)

DT 01-MAY-2000 (TRENDEL, 13, Last sequence update)

DT 01-MAY-2000 (TRENDEL, 13, Last annotation update)

DS TRANSLATION ELONGATION FACTOR (FRAGMENT).

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;

OC Magnoliophyta; Liliopsida; Poales; Poaceae; Oryza.

OX NCBI_TaxID:4530;

XP SEQUENCE FROM N.A.

RA "Inducible expression of translation elongation factor 1A gene in rice

RL seedlings in response to environmental stresses.";

RL Acta Bot. Sin. 41:800-806(1999).

DR EMBL; AF067195; AAC79991.1; .

KW Elongation factor.

FT NON_TER 1

SQ SEQUENCE 15 AA: 1514 MW: 955046F069C9775 CRC64;

Query Match 100.0%; Score 22; DB 10; Length 15;

Identical Similarity 100.0%; Pct Id: 28;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KAAAK 5

DB 9 KAAAK 13

|||||

RESULT 2

Q95825 PRELIMINARY; PRT: 24 AA.

ID Q95825

DT 01-MAY-2000 (TRENDEL, 13, Created)

DT 01-MAY-2000 (TRENDEL, 13, Last sequence update)

DT 01-MAY-2000 (TRENDEL, 13, Last annotation update)

DS PHOTOSYSTEM I REACTION CENTER SUBUNIT IV, PSI-E.

OS Spinacia oleracea (Spinach).

CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliopsida; eudicotyledons; core eudicots; Caryophyllidae;
 CC OX NCBI_TaxID=3562;
 CC (1)
 RN SEQUENCE.
 RX MEDLINE=9249324; PubMed=1374333;
 RX Lagoutte B., Vallon O.;
 RT "Purification and membrane topology of PSI-D and PSI-E, two subunits
 of the photosystem I reaction center.";
 RL Eur. J. Biochem. 205:1175-1185(1992).
 SQ SEQUENCE 24 AA; 2150 MW; 5CD2898B5B18955B CRC64;

Query Match 100.0%; Score 22; DB 10; Length 24;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
 Db 20 KAAK 24
 |||||

RESULT 3
 ID Q01215 PRELIMINARY; PRT; 45 AA.
 DT 01-NOV-1995 (TRENBLREL. 01, Created)
 DT 01-NOV-1995 (TRENBLREL. 09, Last sequence update)
 DT 01-JAN-1999 (TRENBLREL. 12, Last annotation update)
 DE ELONGATION FACTOR 1-ALPHA (PEP-2) (FRAGMENT).
 OS Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
 CC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
 CC Mucor.
 CX NCBI_TaxID=4841;
 RN SEQUENCE FROM N.A.
 RP "The histone H1 complement of *Casnorhabditis elegans*.";
 RX Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RX "The histone H1 complement of *Casnorhabditis elegans*.";
 RX EXPRESSION: 3 genes for elongation factor 1 alpha during
 RT morphogenesis of *Mucor racemosus*.
 RL Mol. Cell. Biol. 7:1925-1932(1987).
 DR EMBL; M16353; AAA33426.1;
 DR EMBL; M16354; AAA33428.1;
 DR EMBL; 20745; Rhira:3109;20745.
 KW Elongation factor.
 FT NON_TER 1
 SQ SEQUENCE 45 AA; 4801 MW; C3E9BD03229AD80 CRC64;

Query Match 100.0%; Score 22; DB 3; Length 45;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
 Db 37 KAAK 41
 |||||

RESULT 4
 ID Q03005 PRELIMINARY; PRT; 58 AA.
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
 DE S21 RIBOSOMAL PROTEIN.
 GN RS21.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 CX NCBI_TaxID=83558;
 RN SEQUENCE FROM N.A.

CC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871562;
 RX Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RX "Characterization of the histone H1 complement of *Chlamydia pneumoniae* J138
 from Japan and OHL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 DR EMBL; AF002545; BAA98243.1;
 KW Ribosomal protein.
 SQ SEQUENCE 58 AA; 6636 MW; 74118AE525600177 CRC64;

Query Match 100.0%; Score 22; DB 2; Length 58;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
 Db 50 KAAK 54
 |||||

RESULT 5
 ID Q093K3 PRELIMINARY; PRT; 60 AA.
 DT 01-MAY-2000 (TRENBLREL. 13, Created)
 DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
 DE HISTONE H1.
 OS *Casnorhabditis elegans*.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditidae;
 CC Rhabditidae; Peloderinae; *Casnorhabditis*.
 CX NCBI_TaxID=6239;
 RN SEQUENCE FROM N.A.
 RP Jedrusik M.A., Schulze E.;
 RX Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RX EXPRESSION: 1 gene for histone H1.
 DR EMBL; M16353; AAA33426.1;
 DR EMBL; 20745; Rhira:3109;20745.
 KW Elongation factor.
 FT NON_TER 1
 SQ SEQUENCE 60 AA; 6212 MW; D01ABBA4CEC3556D CRC64;

Query Match 100.0%; Score 22; DB 5; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
 Db 44 KAAK 48
 |||||

RESULT 6
 ID Q09EJ7 PRELIMINARY; PRT; 61 AA.
 DT 01-OCT-2000 (TRENBLREL. 15, Created)
 DT 01-OCT-2000 (TRENBLREL. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLREL. 15, Last annotation update)
 DE HISTONE H1.
 GN H1A61.
 OS *Trypanosoma brucei gambiense*.
 CC Eukaryota; Eukaryota; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 CX NCBI_TaxID=31255;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-STIB 755;
 RA Grueter E.;
 RL Thesis (2000), Department of Parasitology, Institute of Zoology,
 RL SWITZERLAND.
 DR EMBL; A1287597; CAB76179.1;
 SQ SEQUENCE 61 AA; 6059 MW; F60CE6D6D6A73F CRC64;

Query Match 100.0%; Score 22; DB 5; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 55 KAAK 59
 ID Q9NFJ5 PRELIMINARY; PRT; 61 AA.
 AC Q9NFJ5;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE HISTONE H1.
 RC STRAIN-TRU 927/4;
 RP SEQUENCE FROM N.A.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5702;
 RN [1]
 RA Grueter E.;
 RL Thesis (2000), Department of Parasitology, Institute of Zoology,
 RL SWITZERLAND.
 DR EMBL; AJ287603; CAB76188.1; -
 SQ SEQUENCE 61 AA; 5012 MW; 7B65DA10A7D33893 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 61;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 55 KAAK 59
 ID Q17536 PRELIMINARY; PRT; 62 AA.
 AC Q17536;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE COB10.5 PROTEIN.
 GN COB10.5.
 OS Caenorhabditis elegans.
 OC Metazoa; Chordata; Chordata; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermatidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RA Wilson R.;
 RL MEDLINE=94150718; PubMed=7906398;
 RC STRAIN-BRISTOL NZ;
 RA Bonfield J., Burton J., Connell M., Copsey T., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Graden A., Green S., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Karsenti G., Karsenti G., Karsenti G., Karsenti G.,
 RA Lagtning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shonkhen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Wainstock L., Wilkinson-Sproat J., Wohlman P.;
 RL "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL Nature 368:32-38(1994).
 RN [2]

Query Match 100.0%; Score 22; DB 5; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 55 KAAK 59
 ID Q17536 PRELIMINARY; PRT; 62 AA.
 AC Q17536;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE COB10.5 PROTEIN.
 GN COB10.5.
 OS Caenorhabditis elegans.
 OC Metazoa; Chordata; Chordata; Rhabditida; Rhabditidae;
 OC Rhabditidae; Pelodermatidae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RA Wilson R.;
 RL MEDLINE=94150718; PubMed=7906398;
 RC STRAIN-BRISTOL NZ;
 RA Bonfield J., Burton J., Connell M., Copsey T., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Graden A., Green S., Hawkins T., Hillier L., Jier M., Johnston L.,
 RA Jones M., Karsenti G., Karsenti G., Karsenti G., Karsenti G.,
 RA Lagtning J., Lloyd C., Murray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shonkhen R.,
 RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Wainstock L., Wilkinson-Sproat J., Wohlman P.;
 RL "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans";
 RL Nature 368:32-38(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL NZ;
 RA Blanchard M., Bradshaw H.;
 RL Submitted (May-1995) to the EMBL/GenBank/DBAJ databases.
 DR EMBL; U58757; AAC47916.1; -
 SQ SEQUENCE 62 AA; 6453 MW; FF301204D637D406 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 62;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 46 KAAK 50
 ID Q9NFJ3 PRELIMINARY; PRT; 66 AA.
 AC Q9NFJ3;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE HISTONE H1.
 GN H1A66.
 OS Trypanosoma brucei brucei.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5702;
 RN [1]
 RA Grueter E.;
 RL Thesis (2000), Department of Parasitology, Institute of Zoology,
 RL SWITZERLAND.
 DR EMBL; AJ287608; CAB76193.1; -
 SQ SEQUENCE 66 AA; 6593 MW; 2D664C3471064DC6 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 66;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KAAK 5
 DB 60 KAAK 64
 ID Q9NFJ8 PRELIMINARY; PRT; 71 AA.
 AC Q9NFJ8;
 DT 01-OCT-2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE HISTONE H1.
 GN H1A71.
 OS Trypanosoma brucei gambiense.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=31285;
 RN [1]
 RA Grueter E.;
 RL Thesis (2000), Department of Parasitology, Institute of Zoology,
 RL SWITZERLAND.
 DR EMBL; AJ287594; CAB76176.1; -
 SQ SEQUENCE 71 AA; 7019 MW; 6C7DED4F3E75D8C4 CRC64;

Query Match 100.0%; Score 22; DB 5; Length 71;
 Best Local Similarity 100.0%; Pred. No. 1.e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KAAK 5
|||||
Db 65 KAAK 69

Qy 1 KAAAK 5
11111
Db 69 KAAAK 73

RESULT 15
Q9Y5V0
ID Q9Y5V0 PRELIMINARY; PRT; 76 AA.
AC Q9Y5V0;
DT 01-NOV-1999 (TRENDELrel. 12, Created)

DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DS HSPC038 PROTEIN.
OS Homo sapiens (Human).
OC Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ye M., Zhang Q., Zhou J., Shen Y., Guan Z., Wu X., Fan H., Mao H.,
RA Dai M., Huang Q., Chen S., Chen Z.;
RT "Human HSPC038 mRNA, complete cds."
RL Submitted (1999-1393) to the EMBL/GenBank/DBJ databases.
DI AB000000.1 (1999-1393) 161.
DR INTERPRO: IPR000822.
DR PFM: PF00056; zf-C2H2: 1.
DR PROSITE: PS00028: ZINC_FINGER_C2H2: 1.
DR zinc-finger: Metal-binding: DNA-binding.
SEQUENCE 76 AA: 8498 MW: 3EDFAD9FE2529FF CRC64:

Query Match 100.0%; Score 22; DS 4; Length 76;
Best local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KAAK 5
DB 30 KAAK 34

Search completed: April 24, 2001, 16:40:32
JOB time: 422 sec

GenScan version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:38:18 ; Search time 115.25 seconds

(without alignments)
2.976 Million cell updates/sec

Title: US-09-340-736-5
Perfect score: 31
Sequence: 1 POGVA 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Scheduled: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

A_Genescq_0401.*

1: /SIDS1/gcgdata/genescq/AA1980.DAT.*

2: /SIDS1/gcgdata/genescq/AA1981.DAT.*

3: /SIDS1/gcgdata/genescq/AA1982.DAT.*

4: /SIDS1/gcgdata/genescq/AA1983.DAT.*

5: /SIDS1/gcgdata/genescq/AA1984.DAT.*

6: /SIDS1/gcgdata/genescq/AA1985.DAT.*

7: /SIDS1/gcgdata/genescq/AA1986.DAT.*

8: /SIDS1/gcgdata/genescq/AA1987.DAT.*

9: /SIDS1/gcgdata/genescq/AA1988.DAT.*

10: /SIDS1/gcgdata/genescq/AA1989.DAT.*

11: /SIDS1/gcgdata/genescq/AA1990.DAT.*

12: /SIDS1/gcgdata/genescq/AA1991.DAT.*

13: /SIDS1/gcgdata/genescq/AA1992.DAT.*

14: /SIDS1/gcgdata/genescq/AA1993.DAT.*

15: /SIDS1/gcgdata/genescq/AA1994.DAT.*

16: /SIDS1/gcgdata/genescq/AA1995.DAT.*

17: /SIDS1/gcgdata/genescq/AA1996.DAT.*

18: /SIDS1/gcgdata/genescq/AA1997.DAT.*

19: /SIDS1/gcgdata/genescq/AA1998.DAT.*

20: /SIDS1/gcgdata/genescq/AA1999.DAT.*

21: /SIDS1/gcgdata/genescq/AA2000.DAT.*

22: /SIDS1/gcgdata/genescq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	31	100.0	16	P26313	Sequence of beta-t
2	31	100.0	18	P26382	Sequence of beta-t
3	31	100.0	18	R39599	Scintigraph image
4	31	100.0	18	W11072	Atherosclerotic pl
5	31	100.0	18	W31122	Atherosclerotic pl
6	31	100.0	19	R15135	Vascular injury af
7	31	100.0	19	R42542	Leukocyte-binding
8	31	100.0	19	W47337	Inhibitor of LDL b
9	31	100.0	19	Y55876	Apolipoprotein fra
10	31	100.0	19	Y31101	Elastin derived pe
11	31	100.0	20	R40016	Scintigraph image

12	31	100.0	20	14	R42535	Leukocyte-binding
13	31	100.0	21	14	R42534	Leukocyte-binding
14	31	100.0	21	14	R42534	Leukocyte-binding
15	31	100.0	22	7	P61339	Sequence of chemot
16	31	100.0	22	12	R14949	Medial middle laye
17	31	100.0	24	14	R42538	Leukocyte-binding
18	31	100.0	28	12	R14979	Part of elastomeri
19	31	100.0	28	12	R14986	Part of elastomeri
20	31	100.0	28	12	R15021	Part of elastomeri
21	31	100.0	28	12	R15028	Part of elastomeri
22	31	100.0	30	12	R14953	Part of elastomeri
23	31	100.0	30	12	R14953	Part of elastomeri
24	31	100.0	30	12	R14953	Part of elastomeri
25	31	100.0	30	12	R15000	Part of elastomeri
26	31	100.0	30	12	R14965	Part of elastomeri
27	31	100.0	30	12	R15014	Part of elastomeri
28	31	100.0	30	12	R14972	Part of elastomeri
29	31	100.0	30	12	R15007	Part of elastomeri
30	31	100.0	42	21	R98629	Peptide containing
31	31	100.0	112	20	Y74244	Human prostate tum
32	31	100.0	114	7	P60728	Synthetic elastome
33	31	100.0	114	7	P60728	Synthetic elastome
34	31	100.0	203	19	R46316	Human tropoelastin
35	31	100.0	203	19	R46316	Human tropoelastin
36	31	100.0	515	21	Y69135	Amino acid sequenc
37	31	100.0	571	21	Y69071	Truncated plant-op
38	31	100.0	605	20	Y16791	Human tropoelastin
39	31	100.0	660	20	Y01303	Human tropoelastin
40	31	100.0	698	20	Y01302	Human tropoelastin
41	31	100.0	698	20	Y01302	Amino acid sequenc
42	31	100.0	712	21	R98630	Amino acid sequenc
43	31	100.0	712	21	R98630	Amino acid sequenc
44	31	100.0	730	19	R46314	Human tropoelastin
45	31	100.0	730	19	R46314	Human tropoelastin
	31	100.0	730	21	R98631	Fusion protein com
	31	100.0	731	21	Y69068	Amino acid sequenc

ALIGNMENTS

RESULT 1

P91313
P91313 standard; Protein: 16 AA.

XX P91313;
XX P91313;

AC P91313;
AC P91313;

DT 09-MAR-1992 (first entry)
DT 09-MAR-1992 (first entry)

DE Sequence of beta-turn of a bioelastomeric material.
DE Sequence of beta-turn of a bioelastomeric material.

XX Bioelastomer: elastomeric material; wound dressing; burn;
XX Bioelastomer: elastomeric material; wound dressing; burn;

XX Artificial veins; arteries; skin; ligament; biodegradable.
XX Artificial veins; arteries; skin; ligament; biodegradable.

XX Key Location/Qualifiers
XX Key Location/Qualifiers

FT Misc-difference 1..5
FT Misc-difference 1..5

FT Misc-difference 12..16
FT Misc-difference 12..16

FT Misc-difference 6..11
FT Misc-difference 6..11

FT Misc-difference 1..16
FT Misc-difference 1..16

XX W08910095-A.
XX W08910095-A.

XX 02-NOV-1989.
XX 02-NOV-1989.

XX 14-APR-1989; 89WO-0501482.
XX 14-APR-1989; 89WO-0501482.

XX 21-APR-1989; 89US-0184407.
XX 21-APR-1989; 89US-0184407.

XX (UBER-) URB-RES FOUNDATION.
XX (UBER-) URB-RES FOUNDATION.

XX Urry DW;
 XX NPI; 1989-339743/46.
 XX
 XX Elastomeric polypeptide material - a useful for preventing
 XX adhesion between tissues and wound repair sites
 XX
 XX Claim 13; page 89; 91pp; English.
 XX
 XX The elastomeric material of the invention comprises a bioelastomer
 XX contg. repeating elastoeic tetrapeptide or pentapeptide units opt.
 XX modified by hexapeptide units, the units consisting of hydrophobic
 XX AA residues and gly residues and existing in a conformation having
 XX a polypeptide unit vector (see P91307 and P91308) and comprising
 XX a polypeptide repeating unit of formula VPXG (see P91308), or a
 XX pentapeptide repeating unit of formula VPXG (see P91309),
 XX IFEG (see P91310), or XFXG (see P91312); and the hexapeptide
 XX repeating unit of formula ARGVGV (see P91313).
 XX
 XX Sequence 16 AA;
 XX
 XX Query Match 100.0%; Score 31; DB 10; Length 16;
 XX Best Local Similarity 100.0%; Pred. No. 11;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 PGVGA 6
 XX Db 1 P9VGA 6
 XX
 XX RESULT 2
 XX R26382
 XX ID R26382 standard; peptide; 18 AA.
 XX AC R26382;
 XX DT 25-JAN-1993 (first entry)
 XX DE Sequence of peptides targeted to infections and atherosclerotic
 XX plaque.
 XX KW Technetium-99m labelled polypeptide imaging agent;
 XX radiolabelled imaging; radiodiagnostic agent.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Peptide 1..6
 XX Note= "claimed"
 XX
 XX W09213572-A.
 XX PD 20-AUG-1992.
 XX PF 07-FEB-1992; 92WO-US00757.
 XX PR 08-FEB-1991; 91US-0653012.
 XX (DIAT-) DIATECH INC.
 XX Dean RT;
 XX NPI; 1992-299767/36.
 XX
 XX New technetium-99m labelled polypeptide imaging agents - for
 XX imaging of clots, tumours, infection sites, atherosclerotic and
 XX amyloid plaques or bone, and for visualising organs
 XX Claim 6; Page 13; 19pp; English.
 XX

CC The binding peptide is covalently linked to a 'Cp(aa)' technetium
 CC binding group wherein Cp is a protected cysteine and (aa) is an amino
 CC acid. The technetium-99m complexes are used to image target sites
 CC within a mammalian body.
 XX
 XX Sequence 18 AA;
 XX
 XX Query Match 100.0%; Score 31; DB 13; Length 18;
 XX Best Local Similarity 100.0%; Pred. No. 12;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 PGVGA 6
 XX Db 5 PGVGA 10
 XX
 XX RESULT 3
 XX R39999
 XX ID R39999 standard; peptide; 18 AA.
 XX AC R39999;
 XX DT 23-MAY-1994 (first entry)
 XX DE Scintigraph imaging agent specific binding peptide.
 XX KW Reagent; site imaging; technetium-99m labelled peptide.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Modified-site 18
 XX Note= "C-terminal amide"
 XX
 XX W09321962-A.
 XX PD 11-NOV-1993.
 XX PF 19-APR-1993; 93WO-US03607.
 XX PR 30-APR-1992; 92US-0871282.
 XX (DIAT-) DIATECH INC.
 XX Buttram S, Civitello ER, Dean RT, Lister-James J, McBride W;
 XX NPI; 1993-368425/46.
 XX
 XX Reagents for preparing scintigraphic imaging agents - contg.
 XX technetium-99m labelled peptide(s) contg. 3-100 aminoacid(s)
 XX Claim 35; Page 39; 55pp; English.
 XX
 XX The sequence is that of a specific binding peptide used as part of
 XX a reagent for preparing a scintigraphic imaging agent for imaging
 XX sites within a mammalian body. In this the peptide is covalently
 XX linked to a radiolabel-binding moiety which is capable of forming
 XX a complex with a radioisotope, pref. technetium-99m.
 XX
 XX Sequence 18 AA;
 XX
 XX Query Match 100.0%; Score 31; DB 14; Length 18;
 XX Best Local Similarity 100.0%; Pred. No. 12;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 PGVGA 6
 XX Db 5 PGVGA 10
 XX
 XX RESULT 4

WI1072
 ID WI1072 standard; peptide: 18 AA.
 XX
 XX
 XX WI1072;
 DT 03-JUN-1997 (first entry)
 XX
 XX Atherosclerotic plaque targeting peptide used in diagnostic imaging.
 DE
 DE Leukocyte target; direct; chelator; radionuclide; radiolabel;
 KW isotope; atherosclerosis; thrombosis; embolism; infection; thrombus;
 KW diagnosis; imaging.
 XX
 XX Synthetic.
 XX
 XX WO9603427-A1.
 PN
 XX 08-FEB-1996.
 XX
 XX 28-APR-1995; 95WO-CR00249.
 XX
 XX 22-JUL-1994; 94US-C079155.
 XX
 XX (RESO-) RESOLUTION PHARM INC.
 XX
 XX Goodbody A, Pollak A;
 PI
 XX WPI; 1996-116994/12.
 DR
 XX New peptide derived radionuclide chelators and metal complexes
 PT useful for diagnostic imaging
 PT
 XX
 XX Disclosure; Page 7; 30pp; English.
 XX
 CC WI1072-WI1086 are peptides used for targeting agents to an
 CC atherosclerotic plaque or site of infection (no further details are
 CC given in the specification). The peptides may be coupled to a chelator
 CC molecule, which is labelled with a diagnostically useful metal nuclide
 CC to form a peptide derived radionuclide chelator molecules. Such
 CC targeted and labelled chelators are used to detect pathological
 CC conditions by diagnostic imaging. Radionuclides used include 99mTc,
 CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 188Re, 188Re, 198Au, 203Pb, 212Pb
 CC and 212Bi. The coupling of a targeting agent and radionuclide using a
 CC chelating agent is an alternative to the direct labelling of targeting
 CC agent with a radionuclide. The direct labelling of targeting agent
 CC with a radionuclide is also possible. The new conjugates give
 CC low-affinity sites, forming unstable complexes. The new conjugates give
 CC better scintigraphic images in rat inflammation studies than known
 CC imaging agents Ga-67, 99mTc-196, 111In-WC and 99mTc-Nancocoll. They image
 CC more rapidly than the known agents and show superior biodistribution.
 XX
 XX Sequence 18 AA;
 SQ
 Query Match 100.0%; Score 31; DB 17; Length 18;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 5 pgvga 10
 RESULT 5
 WI1122
 ID WI1122 standard; peptide: 18 AA.
 XX
 XX WI1122;
 DT 23-JAN-1998 (first entry)
 XX
 XX Atherosclerotic plaque- and infection site-targeting peptide.
 DE
 XX Target; delivery; radionuclide chelator; diagnosis; therapy;
 KW

XX detection; atherosclerosis; thrombosis; platelet.
 XX Synthetic.
 XX US5659041-A.
 PN
 XX 19-AUG-1997.
 PD
 XX 19-JUL-1993; 93US-0092911.
 XX
 XX 02-SEP-1994; 94US-0299636.
 PR 19-JUL-1993; 93US-0092911.
 XX
 XX (RESO-) RESOLUTION PHARM INC.
 XX
 XX Dunn-Dufault R, Kirby RA, Pollak A;
 PI
 XX WPI; 1997-424290/39.
 DR
 XX New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
 PT chelator of radionuclides for radio:imaging of target tissues of
 PT diagnostic interest
 PT
 XX Disclosure; Column 17-18; 20pp; English.
 XX
 CC W31110-W31147 are peptides used for targeting a new hydrazino-type
 CC compound to various sites of disease, e.g. atherosclerotic plaque,
 CC sites of infection, platelets, thrombus or amyloid plaque. The new
 CC compound is a radionuclide chelator and is used to radiolabel the
 CC targeting peptides for the detection and diagnostic imaging of
 CC sites of disease, e.g. amyloid plaques in Alzheimer's disease.
 CC
 XX Sequence 18 AA;
 SQ
 Query Match 100.0%; Score 31; DB 18; Length 18;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 5 pgvga 10
 RESULT 5
 RI5135
 ID RI5135 standard; Protein; 19 AA.
 XX
 XX RI5135;
 AC
 XX 18-FEB-1992 (first entry)
 DT
 XX Vascular injury affinity peptide.
 DE
 XX Low density lipoprotein; atherosclerosis.
 KW
 XX Synthetic.
 OS
 XX WO9116919-A.
 FN
 XX 14-NOV-1991.
 PD
 XX 02-MAY-1991; 91WO-US03026.
 PP
 XX 03-MAY-1990; 90US-0518215.
 PR 03-MAY-1990; 90US-0518142.
 XX
 XX (NEMF-) NEW ENGLAND DEACON.
 PA
 XX Lees RS, Lees AM, Fischman A, Shih IL, Findeis MA;
 PI
 XX WPI; 1991-353525/48.
 DR
 XX

PT Synthetic peptide(s) comprising amphiphilic domain of apoA-I -
 PT used to diagnose vascular injury or disease or inhibit binding of
 PT low density lipoprotein to vascular walls in treating
 PT atherosclerosis
 XX Claim 25; Page 48; 66pp; English.
 XX
 CC The amino acid sequence is that of a synthetic peptide derived from
 CC elastin, a vascular associated protein. The peptide, which may opt.
 CC be labelled, is used to detect injuries in the vascular system, esp.
 CC atherosclerosis in its early stages before it causes stenosis and
 CC blood flow disturbances. It can also be used to inhibit binding of
 CC low density lipoprotein (LDL) to vascular walls, i.e. to prevent
 CC or alleviate atherosclerosis. It is easy to prepare on a large scale
 CC and allows vascular regions to be located non-invasively without
 CC complex equipment or highly skilled personnel. See also R15126-R15140.
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 12; Length 19;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 6 P9VGA 11
 |||||

RESULT 7
 R42542 ID R42542 standard; Peptide: 19 AA.
 XX AC R42542;
 XX DT 05-AUG-1994 (first entry)
 DE Leukocyte-binding peptide which can bind to technetium-99m.
 XX Leukocyte binding peptide; elastin; scintigraphic imaging;
 KW Inflammation site; technetium 99m.
 OS Synthetic.

Key Location/Qualifiers
 Modified-site 1
 /label= OTHER
 /note= "BMT, i.e. M6/N9-bis(2-methyl-
 2-mercaptopropyl)-6,9-diazanonoic acid"
 Modified-site 19
 /note= "Gly-NH2"
 XX W09317719-A.
 XX
 PD 16-SEP-1993.
 XX
 PF 12-MAR-1993; 93WO-US02320.
 XX
 PR 13-MAR-1992; 92US-0851074.
 XX
 PA (DIAT-) DIATECH INC.
 XI Buttram S, Dean RT, Lees RS, Lister-James J;
 DR WPI; 1991-303154/38.
 XX
 PT Scintigraphic imaging agent for sites of inflammation -
 PT comprising leukocyte-binding peptide bound technetium-99m via
 PT binding moiety
 XX Claim 14; Page 31; 40pp; English.
 XX
 CC New peptides are claimed which are leukocyte binding peptides

CC Having covalently bound to them a moiety which can bind a Tc-99m
 CC radiolabel. The peptides having Tc-99m bound to them are useful as
 CC scintigraphic imaging agents for imaging sites of infection and
 CC inflammation in the mammalian body, e.g. caused by ischaemia,
 CC inflammatory bowel disorder, arthritis or tumours.
 CC The present sequence is a specifically claimed example of
 CC such a peptide derived from elastin.
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 14; Length 19;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 6 P9VGA 11
 |||||

RESULT 8
 W47337 ID W47337 standard; peptide: 19 AA.
 XX AC W47337;
 XX DT 01-JUN-1998 (first entry)
 DE Inhibitor of LDL binding to vascular wall.
 XX Binding inhibitor; low-density lipoprotein; LDL; vascular wall;
 KW vascular injury; elastin; collagen; prevention; treatment;
 KW vascular disease; atherosclerosis.
 OS Synthetic.

US5726153-A.
 XX
 PD 10-MAR-1998.
 XX
 PF 06-JUN-1995; 95US-0468543.
 PR 06-JUN-1995; 95US-0468543.
 PR 02-MAY-1988; 88US-0183130.
 PR 03-MAY-1990; 90US-0518142.
 PR 03-MAY-1990; 90US-0518215.
 PR 02-MAY-1991; 91US-0694929.
 PR 16-APR-1993; 93US-0048569.
 PR 24-FEB-1994; 94US-0201057.
 PR 28-FEB-1995; 95US-0398046.
 PR 06-JUN-1995; 95US-0468543.
 XX
 XX (NEW-) NEW ENGLAND DEACONESS HOSPITAL.
 XI Finkel MA, Fischman A, Lees RM, Lees RS, Shih I;
 DR WPI; 1998-192802/17.
 XX
 PT Inhibiting binding of low-density lipoprotein to vascular walls - by
 PT administering peptide with affinity for vascular injury sites

PS Disclosure; Column 5; 3pp; English.
 XX
 CC A novel method for inhibiting the binding of a low-density
 CC lipoprotein (LDL) to vascular walls in vivo, comprises
 CC administering to a patient a peptide containing an
 CC amphiphilic domain and having affinity for sites of vascular
 CC injury, e.g. the present peptide.
 CC The peptide inhibits the binding of LDL to vascular wall
 CC components, e.g. elastin and collagen, and so can be used to
 CC prevent or treat vascular diseases, e.g. atherosclerosis.
 XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 19; Length 19;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 6 pgvga 11
 |||||

RESULT 9
 ID 155876
 ID 155876 standard; peptide: 19 AA.
 AC Y55876;
 XX
 XX 01-FEB-2000 (first entry)

Apolipoprotein fragment peptide #26 for vascular disease imaging.
 XX
 KW Diagnosis; water soluble; amphiphilic domain; affinity; vascular injury;
 KW detection; imaging; disease; atherosclerosis; apolipoprotein.
 XX
 OS Synthetic.
 OS Homo sapiens.

PN US5572890-A.
 XX
 XX 26-OCT-1999.
 PD
 XX 28-FEB-1995;
 PF 95US-0398046.
 XX

XX 02-MAY-1991;
 PR 91US-0694929.
 PR 16-APR-1993;
 PR 93US-0048559.
 PR 24-FEB-1994;
 PR 94US-0201057.
 PR 02-MAY-1988;
 PR 88US-0181130.
 PR 03-MAY-1990;
 PR 90US-0318142.
 XX 03-MAY-1990;
 XX 90US-0518215.
 PA (NEMF-) NEW ENGLAND DEACONESS HOSPITAL.

XX
 XX Lees AM, Fischman A, Shih I, Vindeis WA, Lees RS;
 XX WPI; 1999-632541/54.

New diagnostic synthetic peptides which have affinity for and
 accumulate at a site of vascular injury useful for detection and
 imaging of vascular disease such as atherosclerosis
 XX
 PS Disclosure; Column 5; 30pp; English.

XX The peptides Y55851-Y55889 represent examples of diagnostic, synthetic
 CC peptides which carry a detectable label, contain 30 or fewer amino acids,
 CC are water soluble, contain an amphiphilic domain and have affinity for,
 CC and propensity to accumulate at, a site of vascular injury. They are
 CC preferably derived from the amino acid sequence of apolipoprotein. The
 CC peptides can be used for the detection or imaging of a vascular injury
 CC or disease, e.g. atherosclerosis.

XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 20; Length 19;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 6 pgvga 11
 |||||

RESULT 10

Y33101
 ID Y33101 standard; peptide: 19 AA.
 XX
 AC Y33101;
 XX
 XX 15-NOV-1999 (first entry)
 YZ
 DE Elastin derived peptide 4.
 XX
 KW Apolipoprotein B; apolipoprotein AI; elastin; vascular imaging; detection;
 KW injury; vascular system; atherosclerosis; re-endothelializing; noninvasive;
 KW aortic lesion; trauma; lipoprotein accumulation.

XX
 OS Synthetic.

XX US5955055-A.

XX 21-SEP-1999.

XX 06-JUN-1995;
 PF 95US-0469592.

XX 02-MAY-1991;
 PR 91US-0694929.

PR 02-MAY-1988;
 PR 88US-0181130.

PR 03-MAY-1990;
 PR 90US-0518142.

PR 16-APR-1993;
 PR 93US-0048559.

PR 24-FEB-1994;
 PR 94US-0201057.

PR 02-FEB-1995;
 PR 95US-0398046.

XX 06-JUN-1995;
 XX 95US-0469592.

PA (NEMF-) NEW ENGLAND DEACONESS HOSPITAL.

XX Vindeis WA, Fischman A, Lees AM, Lees RS, Shih I;

XX WPI; 1999-539543/45.

PI Detecting vascular injuries using a labeled peptide useful for the

CC diagnosis and monitoring of atherosclerosis

XX
 PS Disclosure; Column 27-28; 31pp; English.

XX This invention describes a novel method (I) for detecting injuries in a

CC vascular system using labeled synthetic peptides which have affinity for

CC and propensity to accumulate at a site of vascular injury. The method is

CC particularly suitable for detecting and monitoring atherosclerosis. It

CC has been found that the synthetic peptide accumulates at the healing

CC (re-endothelializing) edge of aortic lesions produced by trauma. These

CC lesions resemble human atherosclerosis in many important respects

CC (including accumulation of lipoproteins and other pathological changes).

CC The ability of the synthetic peptides to localize at the trauma site, and

CC to permit imaging, may be used to visualize vascular disease. The method

CC is noninvasive and the peptides used do not target vascular lesions. The

CC peptides are water soluble, contain an amphiphilic domain and have affinity for

CC and propensity to accumulate at a site of vascular injury. They are

CC represent synthetic peptides derived from apolipoprotein B, apolipoprotein AI

XX and elastin which are used in the method of the invention.

XX
 SQ Sequence 19 AA;

Query Match 100.0%; Score 31; DB 20; Length 19;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6

DB 6 pgvga 11

|||||

RESULT 11

R40018

ID R40018 standard; peptide: 20 AA.

XX

AC R40018;

XX 23-MAY-1994 (first entry)
 DT Scintigraph imaging agent specific binding peptide.
 XX Reagent; site imaging; technetium-99m labelled; peptide.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 20 /note= "C-terminal amide"
 FT
 XX WO9321962-A.
 XX 11-NOV-1993.
 XX 19-APR-1993; 93WO-US03687.
 XX 30-APR-1992; 92US-0871282.
 XX (DIAT-) DIATECH INC.
 XX Buttram S, Civitello ER, Dean RT, Lister-James J, McBride W;
 XX WPI; 1993-368429/46.
 XX Reagents for preparing scintigraphic imaging agents - contg.
 FT technetium-99m labelled peptide(s) contg. 3-100 aminoacid(s)
 XX Claim 35; Page 39; 55pp; English.
 XX The sequence is that of a specific binding peptide used as part of
 CC a reagent for preparing a scintigraphic imaging agent for imaging
 CC sites within a mammalian body. In this the peptide is covalently
 CC linked to a radiolabel-binding moiety which is capable of forming
 CC a complex with a radioisotope, pref. technetium-99m.
 XX Sequence 20 AA;

Query Match 100.0%; Score 31; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PGVGA 6
 Db 7 Pygva 12

Query 12

XX R42535 standard; Peptide; 20 AA.
 XX R42535;
 XX 05-AUG-1994 (first entry)
 XX Leukocyte-binding peptide which can bind to technetium-99m.
 DE Leukocyte binding peptide; elastin; scintigraphic imaging;
 KW inflammation site; technetium 99m.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "Picolinoyl-Gly"
 FT Modified-site 2 /note= "acetamidomethyl-Cys"
 FT Modified-site 20 /note= "Gly-NH2"

XX WO9317719-A.
 XX 16-SEP-1993.
 XX 12-MAR-1993; 93WO-US02320.
 XX 13-MAR-1992; 92US-0851074.
 XX (DIAT-) DIATECH INC.
 XX Buttram S, Dean RT, Lees RS, Lister-James J;
 FT

PN WO9317719-A.
 XX 16-SEP-1993.
 XX 12-MAR-1993; 93WO-US02320.
 XX 13-MAR-1992; 92US-0851074.
 XX (DIAT-) DIATECH INC.
 XX Buttram S, Dean RT, Lees RS, Lister-James J;
 XX WPI; 1993-303154/38.
 XX Scintigraphic imaging agent for sites of inflammation -
 FT comprising leukocyte-binding peptide bound technetium-99m via
 FT binding moiety
 XX Claim 14; Page 31; 40pp; English.
 XX New peptides are claimed which are leukocyte binding peptides
 CC having covalently bound to them a moiety which can bind a Tc-99m
 CC radiolabel. The peptides having Tc-99m bound to them are useful as
 CC scintigraphic imaging agents for imaging sites of infection and
 CC inflammation in the mammalian body, e.g. caused by ischaemia,
 CC inflammatory bowel disorder, arthritis or tumours.
 CC The present sequence is a specifically claimed example of
 CC such a peptide derived from elastin.
 XX Sequence 20 AA;

Query Match 100.0%; Score 31; DB 14; Length 20;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PGVGA 6
 Db 7 Pygva 12

RESULT 13
 ID R42541
 XX R42541 standard; Peptide; 20 AA.
 XX R42541;
 XX 05-AUG-1994 (first entry)
 XX Leukocyte-binding peptide which can bind to technetium-99m.
 XX Leukocyte binding peptide; elastin; scintigraphic imaging;
 KW inflammation site; technetium 99m.
 XX Synthetic.
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "Picolinoyl-Gly"
 FT Modified-site 20 /note= "Gly-NH2"

XX WO9317719-A.
 XX 16-SEP-1993.
 XX 12-MAR-1993; 93WO-US02320.
 XX 13-MAR-1992; 92US-0851074.
 XX (DIAT-) DIATECH INC.
 XX Buttram S, Dean RT, Lees RS, Lister-James J;
 FT

XX WPI; 1993-303154/38.
 XX Scintigraphic imaging agent for sites of inflammation -
 CC comprising leukocyte-binding peptide bound technetium-99m via
 CC binding moiety
 CC Claim 14; Page 31; 40pp; English.
 XX New peptides are claimed which are leukocyte binding peptides
 CC which can be used in the treatment of inflammation and a radio-
 CC label. The peptides having 7c-99m bound to them are useful as
 CC scintigraphic imaging agents for imaging sites of infection and
 CC inflammation in the mammalian body, e.g. caused by ischaemia,
 CC inflammatory bowel disorder, arthritis or tumours.
 CC The present sequence is a specifically claimed example of
 CC such a peptide derived from elastin.
 CC Sequence 20 AA;
 XX
 XX Query Match 100.0%; Score 31; DB 14; Length 20;
 XX Best Local Similarity 100.0%; Pred No. 14;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 XX |||||
 DB 7 PGVGA 12
 XX
 XX RESULT 14
 XX R42534
 XX ID R42534 standard; Peptide: 21 AA.
 XX AC R42534;
 XX DT 05-AUG-1994 (first entry)
 XX DE Leukocyte-binding peptide which can bind to technetium-99m.
 XX LE Leukocyte binding peptide; elastin; scintigraphic imaging;
 XX inflammation site; technetium 99m.
 XX OS Synthetic.
 XX
 XX Key Location/Qualifiers
 XX Modified-site 1
 XX /note= "acetamidomethyl-Cys"
 XX Modified-site 3
 XX /note= "acetamidomethyl-Cys"
 XX Modified-site 2
 XX /note= "Gly-NH2"
 XX W09317719-A.
 XX 16-SEP-1993.
 XX 12-MAR-1993; 93WO-US02320.
 XX 13-MAR-1992; 92US-0851074.
 XX (DIAT-) DIAPECE INC.
 XX Buttram S, Dean RT, Lees RS, Lister-James J;
 XX WPI; 1993-303154/38.
 XX
 XX Scintigraphic imaging agent for sites of inflammation -
 CC comprising leukocyte-binding peptide bound technetium-99m via
 CC binding moiety
 CC Claim 14; Page 31; 40pp; English.

CC New peptides are claimed which are leukocyte binding peptides
 CC which can be used in the treatment of inflammation and a radio-
 CC label. The peptides having 7c-99m bound to them are useful as
 CC scintigraphic imaging agents for imaging sites of infection and
 CC inflammation in the mammalian body, e.g. caused by ischaemia,
 CC inflammatory bowel disorder, arthritis or tumours.
 CC The present sequence is a specifically claimed example of
 CC such a peptide derived from elastin.
 CC Sequence 21 AA;
 XX
 XX Query Match 100.0%; Score 31; DB 14; Length 21;
 XX Best Local Similarity 100.0%; Pred No. 14;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 XX |||||
 DB 8 PGVGA 13
 XX
 XX RESULT 15
 XX P61339
 XX ID P61339 standard; peptide; 22 AA.
 XX AC P61339;
 XX DT 30-JUL-1991 (first entry)
 XX DE Sequence of chemotactic peptide which includes a repeat hexapeptide
 XX sequence found in tropoelastin from vascular wall.
 XX EL Elastic fibre-forming fibroblasts; prosthetic device.
 XX Key Location/Qualifiers
 XX Misc-difference 1.4
 XX /note= "residues 1-4,1-3,1-2 or 1 may be absent;
 XX N-terminal is bonded to H or biocompatible
 XX N-terminal gp"
 XX Misc-difference 19..22
 XX /note= "residues 19-22,20-22;21-22 or 22 may be
 XX absent; C-terminal is bonded to OH,OH3 or
 XX biocompatible C-terminal gp; ES- non-toxic
 XX metal ion"
 XX Misc-difference 12..17
 XX /note= "repeat hexapeptide; present 1-100 times"
 XX US4605413-A.
 XX 12-AUG-1986.
 XX 19-SEP-1983; 83US-0533670.
 XX 19-SEP-1983; 83US-0533670.
 XX (UVAL-) UNIV ALABAMA.
 XX Urry DW, Senior RM;
 XX WPI; 1986-232093/35.
 XX
 XX Prosthetic device, e.g. artificial blood vessel or skin - having
 CC chemo-tactic peptide in its surface to enhance invasion of
 CC elastic fibre-forming fibroblasts
 CC Claim 1; column 8; 10pp; English.
 CC The repeat sequence is chemotactic for fibroblasts which synthesise
 CC elastic fibre precursor protein in biological systems.
 CC Sequence 22 AA;
 XX

Query Match 100.0%; Score 31; DB 7; Length 22;
Best Local Similarity 100.0%; Pres. No. 15;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qx 1 POGVA 6
 |||||
Db 1 pvgva 6

Search completed: April 24, 2001, 16:36:19
Job time: 419 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:56:23 ; Search time 62.39 Seconds
(without alignments)
1.847 Million cell updates/sec

Title: US-09-340-736-5

Sequence: 1 POGVA 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Scheduled: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents:AA.*

1: /csn2_5/ptodata/2/aa/58.COMB.pep.*

2: /csn2_5/ptodata/2/aa/58.COMB.pep.*

3: /csn2_5/ptodata/2/aa/6A.COMB.pep.*

4: /csn2_5/ptodata/2/aa/6B.COMB.pep.*

5: /csn2_5/ptodata/2/aa/PCUTS.COMB.pep.*

6: /csn2_5/ptodata/2/aa/PCUTS.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	31	100.0	6	2	US-08-911-364-5	Sequence 5, Appl
2	31	100.0	18	1	US-08-127-351-34	Sequence 34, Appl
3	31	100.0	18	1	US-08-480-567B-34	Sequence 34, Appl
4	31	100.0	18	1	US-08-487-221A-34	Sequence 34, Appl
5	31	100.0	18	1	US-08-480-570-34	Sequence 34, Appl
6	31	100.0	18	1	US-08-273-155-14	Sequence 14, Appl
7	31	100.0	18	1	US-08-478-155-14	Sequence 14, Appl
8	31	100.0	18	1	US-08-464-456-13	Sequence 13, Appl
9	31	100.0	18	1	US-08-464-456-13	Sequence 13, Appl
10	31	100.0	18	1	US-08-703-988A-14	Sequence 14, Appl
11	31	100.0	18	1	US-08-470-152-1	Sequence 1, Appl
12	31	100.0	18	1	US-08-463-052-13	Sequence 13, Appl
13	31	100.0	18	2	US-08-480-551-13	Sequence 13, Appl
14	31	100.0	18	2	US-08-612-842-14	Sequence 14, Appl
15	31	100.0	18	2	US-08-290-853-24	Sequence 24, Appl
16	31	100.0	19	1	US-08-468-543-14	Sequence 14, Appl
17	31	100.0	19	2	US-08-398-892-14	Sequence 14, Appl
18	31	100.0	20	1	US-08-488-135-19	Sequence 19, Appl
19	31	100.0	20	1	US-08-470-152-19	Sequence 19, Appl
20	31	100.0	21	1	US-08-472-535-11	Sequence 11, Appl
21	31	100.0	21	1	US-08-472-535-11	Sequence 11, Appl
22	31	100.0	21	1	US-08-484-774-11	Sequence 11, Appl
23	31	100.0	21	1	US-08-484-774-11	Sequence 11, Appl
24	31	100.0	21	1	US-08-484-774-11	Sequence 11, Appl
25	31	100.0	21	2	US-08-290-853-2	Sequence 2, Appl
26	31	100.0	21	2	US-08-290-853-7	Sequence 7, Appl
27	31	100.0	21	2	US-08-290-853-17	Sequence 17, Appl

Sequence 11, Appl
Sequence 12, Appl
Sequence 15, Appl
Sequence 15, Appl
Sequence 13, Appl
Sequence 15, Appl
Patent No. 5250516
Sequence 2, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 40, Appl
Sequence 23, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-08-911-364-5
; Sequence 5, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION:
; APPLICANT: ROTHSTEIN, Aser
; APPLICANT: KELLY, Fred W.
; ATTORNEY/AGENT INFORMATION:
; TITLE OF INVENTION: SELECTING PEPTIDES MODELED ON HUMAN
; TITLE OF INVENTION: ELASTIN AND OTHER FIBROUS PROTEINS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; COMPUTER PROGRAM FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICANT NUMBER: US/08/911,364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA: US 60/033,552
; FILING DATE: 07-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 041082/0104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-911-364-5

Query Match 100.0% Score 31; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 PEGUYA 6
Db 1 PEGUYA 6

RESULT 2
US-08-127-351-34
; Sequence 34, Application US/08127351
; Patent No. 5449761
; GENERAL INFORMATION:
; APPLICANT: BELINKA JR, BENJAMIN A.
; APPLICANT: COUGHLIN, DANIEL J.
; APPLICANT: ALVAREZ, VERNON L.
; REFERENCE: 4980-004-44
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MAIER &
; ADDRESSEE: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 28-SEP-1993
; CLASSIFICATION: 534
; ATTORNEY/AGENT INFORMATION:
; NAME: Villacorta, Gilberto M.
; REGISTRATION NUMBER: 34,038
; REFERENCE/DOCKET NUMBER: 2654-002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 584-1111
; TELEFAX: (703) 584-1124
; TELEX:
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: Peptide
; US-08-127-351-34

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGUYA 6
Db 5 PEGUYA 10

RESULT 4
US-08-487-221A-34
; Sequence 34, Application US/08487221A
; Patent No. 5593656
; GENERAL INFORMATION:
; APPLICANT: BELINKA JR, BENJAMIN A.
; APPLICANT: COUGHLIN, DANIEL J.
; APPLICANT: ALVAREZ, VERNON L.
; REFERENCE: 4980-004-44
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MAIER &
; ADDRESSEE: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ORLON, SPIVAK, MCCLELLAND, MAIER &
; ADDRESSEE: NEUSTADT, P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
; TITLE OF INVENTION: CONSTRUCTS

```

APPLICATION NUMBER: US 08/127,351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELETYPE: (703) 413-3000
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-487-221A-34

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 5
US-08-480-370-34
Sequence 34, Application US/08480370
Patent No. 5698470
GENERAL INFORMATION:
APPLICANT: BELINDA J. BENJAMIN A.
APPLICANT: COUGHLIN, DANIEL J.
APPLICANT: ALVAREZ, VERNON L.
APPLICANT: WOOD, RICHARD
TITLE OF INVENTION: METAL-BINDING TARGETED POLYPEPTIDE
NUMBER OF SEQUENCES: 56
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER &
ADDRESSEE: NEUSTADT,
ADDRESSEE: C. Jefferson Davis Highway, Suite 400
STATE: Virginia
CITY: Arlington
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 28-SEP-1993
REGISTRATION NUMBER: US/08/480,370
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/127,351
FILING DATE: 28-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Villacorta, Gilberto M.
REGISTRATION NUMBER: 34,038
REFERENCE/DOCKET NUMBER: 4980-004-44
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELETYPE: 248955 OPT 1R
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid

TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-480-370-34

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 6
US-08-299-636-13
Sequence 13, Application US/08299636
Patent No. 5659041
GENERAL INFORMATION:
APPLICANT: POLLAK, Alfred
APPLICANT: KIRBY, Robert A.
APPLICANT: KIRBY, Robert A.
TITLE OF INVENTION: HYDRAINO-TYPE RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 02-SEP-1994
CLASSIFICATION: 534
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/092,911
FILING DATE: 18-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/262/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STEREONUMBER: single
TOPOLOGY: linear
US-08-299-636-13

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 7
US-08-279-155-14
Sequence 14, Application US/08279155

Patent No. 5662885
GENERAL INFORMATION:
APPLICANT: GOTTING, Alfred
INVENTOR: GOTTING, Alfred
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKADO, KARMELESTEIN, MURRAY & ORAM LLP
STREET: 655 Fifteenth Street, N. W., Suite 330 - G
CITY: Washington
STATE: DC
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/279/155
FILING DATE: 22-JUL-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURRAY, Robert B.
FIRM: KARMELESTEIN, MURRAY & ORAM LLP
REFERENCE/DOCKET NUMBER: P0074-4005
TELEPHONE: 202/638-5000
TELEFAX: 202/638-5000
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STANDARDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-279-155-14

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCVGYA 6
DB 5 PCVGYA 10

RESULT 8
US-464-456-13
Sequence 13, Application US/08464456
Patent No. 5681541
GENERAL INFORMATION:
APPLICANT: Dean, Richard T
TITLE OF INVENTION: Technetium-99m Labeled Polypeptides for
TITLE OF INVENTION: Imaging
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 10 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/464/456
FILING DATE: 05-JUN-1995

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 5681541, Kevin E
FIRM: KARMELESTEIN, MURRAY & ORAM LLP
REFERENCE/DOCKET NUMBER: 90.1104-V
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-464-456-13

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PCVGYA 6
DB 5 PCVGYA 10

RESULT 9
US-08-486-135-1
Sequence 1, Application US/08486135
Patent No. 5720934
GENERAL INFORMATION:
APPLICANT: Dean, Richard T
APPLICANT: Buttram, Scott
APPLICANT: McBride, William
APPLICANT: Lister-James, John
APPLICANT: Civitello, Edgar R
TITLE OF INVENTION: Technetium-99m Labeled Peptides for
TITLE OF INVENTION: Imaging
NUMBER OF SEQUENCES: 424
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486/135
FILING DATE: 06-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 5720934, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 92,205-N
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 23-24

OTHER INFORMATION: /label= AMIDE
OTHER INFORMATION: /note= "The carboxyl terminus is an amide."
US-08-486-135-1

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 10
US-08-703-988A-14

Sequence 14; Application US/08703988A

Patent No. 5780006

GENERAL INFORMATION:

APPLICANT: POLLAK, Alfred

APPLICANT: GOODBODY, Anne

TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE

TITLE OF INVENTION: CHELATORS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: MIRADO, MARCELSTEIN, MURRAY & ORAM

STREET: 655 Fifteenth Street, N. W., Suite 330

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20005-5701

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/703,988A

FILING DATE: 28-AUG-1996

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/279,155

FILING DATE: 22-JUL-1994

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: MURRAY, Robert B.

REGISTRATION NUMBER: 22,980

REFERENCE/DOCKET NUMBER: P8074-6011

TELEPHONE: 202/638-5000

TELEFAX: 202/638-4810

INFORMATION FOR SEQ ID NO: 14:

LENGTH: 18 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-703-988A-14

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 11
US-08-470-152-1

Sequence 1; Application US/08470152

Patent No. 5780070

GENERAL INFORMATION:

APPLICANT: Deap, Richard T

APPLICANT: Buttram, Scott

APPLICANT: McBride, William

APPLICANT: Lister-James, John

APPLICANT: Civitello, Edgar R

TITLE OF INVENTION: Technetium-99m Labeled Peptides for

TITLE OF INVENTION: Imaging

NUMBER OF SEQUENCES: 19

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: IL

COUNTRY: US

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,152

FILING DATE: 06-JUN-1995

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 5780070, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 92,205-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 18 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: Peptide

FEATURE:

NAME/KEY: Modified site

LOCATION: 23..24

OTHER INFORMATION: /label= AMIDE

OTHER INFORMATION: /note= "The carboxyl terminus is an amide."

US-08-470-152-1

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 5 PGVGA 10

RESULT 12
US-08-463-052-13

Sequence 13; Application US/08463052

Patent No. 5788560

GENERAL INFORMATION:

APPLICANT: Deap, Richard T

APPLICANT: Buttram, Scott

APPLICANT: McBride, William

APPLICANT: Lister-James, John

APPLICANT: Civitello, Edgar R

TITLE OF INVENTION: Technetium-99m Labeled Polypeptides for

TITLE OF INVENTION: Imaging

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Allegretti, Ltd.

STREET: 10 South Wacker Drive Suite 3000

CITY: Chicago

STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,052
FILING DATE: 05-JUN-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 90,1104-V
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312 715 1000
TELEFAX: 312 715 1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-463-052-13

Query Match 100.0%; Score 31; DB 1; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGVGA 6
DB 5 PEGVGA 10

RESULT 13
US-08-480-551-13
Sequence 13; Application US/08480551
Patent No. 5811394
GENERAL INFORMATION:
APPLICANT: Dean, Richard T
TITLE OF INVENTION: Technetium-99m Labeled Polypeptides for
IMAGING
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: Allegretti & Witcoff, Ltd
STREET: 30 South Wacker Drive Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,551
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/264,176
FILING DATE: 08-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: McDonnell, John J
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 90,1104
TELECOMMUNICATION INFORMATION:

STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,842
FILING DATE: 20-MAR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERMAN, RICHARD J
REGISTRATION NUMBER: 39107
REFERENCE/DOCKET NUMBER: 8012-6002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-638-5000
TELEFAX: 202-638-4810
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-612-842-14

Query Match 100.0%; Score 31; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGVGA 6
DB 5 PEGVGA 10

RESULT 14
US-08-612-842-14
Sequence 14; Application US/08612842
Patent No. 5976495
GENERAL INFORMATION:
APPLICANT: COLLAK, ALFRED
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKALDO, MARCELSTEIN, MURRAY & ORAM
STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
STREET: 330
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,842
FILING DATE: 20-MAR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERMAN, RICHARD J
REGISTRATION NUMBER: 39107
REFERENCE/DOCKET NUMBER: 8012-6002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-638-5000
TELEFAX: 202-638-4810
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-612-842-14

Query Match 100.0%; Score 31; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEGVGA 6
DB 5 PEGVGA 10

```
RESULT 15
US-08-290-853-28
Sequence 28 Application US/08290853
Patent No. 5989519
GENERAL INFORMATION:
APPLICANT: Dean, Richard T
APPLICANT: Buttram, Scott
TITLE OF INVENTION: Technetium-99m Labeled Peptides for
TITLE OF INVENTION: Imaging Inflammation
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, Ltd.
Street: 10 South Wacker Drive, Suite 3000
City: Chicago
State: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290,853
FILING DATE: 11-Oct-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION: Kevin E
NAME/ADDRESS: 5885 Ashland, 32403
REGISTRATION NUMBER: 92,112-H
REFERENCE/DOCKET NUMBER: 92,112-H
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 910-221-5317
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
FEATURE:
MOLECULE TYPE: peptide
NAME/KEY: Modified-site
LOCATION: 1..3
OTHER INFORMATION: /label= BAT
OTHER INFORMATION: /note= "The amino terminal residue is
OTHER INFORMATION: linked to a BAT radiolabel binding moiety."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 18
OTHER INFORMATION: /label= Amide
OTHER INFORMATION: /note= "The carboxyl terminus is modified to an
OTHER INFORMATION: amide"
US-08-290-853-28

Query Match 100.0%; Score 31; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 4,4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGYA 6
Db 5 PGVGYA 10
```

Search completed: April 24, 2001, 16:36:24
Job time: 304 sec

A: Reference number: A24758; MUID: 85305763
 A: Accession: G007762
 A: Molecule type: mRNA
 A: Residues: 655-669, 671-716, 732-770 <Y0>
 C: Superfamily: elastin
 C: Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
 F: 762-765/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 2; Length 770;
 Best Local Similarity 100.0%; Pred. NC. 1.4e+02;
 Matches: 6; Conservative 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 111111
 Db 500 PGVGA 505

RESULT 4
 EX0H
 elastin precursor, long splice form - human
 A: Alternate names: tropoelastin
 A: CDS: 22-Jun-1990 sequence, revision 26-Jul-1996 #test_change 22-Jun-1999
 A: Accession: A32707; A33705; A30524; A53891
 R: Indik, Z.; Yeh, H.; Ornstein-Goldstein, N.; Sheppard, P.; Anderson, N.; Rosenbloom, P. Natl. Acad. Sci. U.S.A. 84, 5680-5684, 1987
 C: Keywords: alternative splicing of human elastin mRNA. Indicated by sequence analysis of
 A: Reference number: A32707; MUID: 87289668
 A: Accession: A32707
 A: Molecule type: mRNA
 A: Residues: 1-2740
 A: Cross-references: GB:J02948
 R: Bashir, M.M.; Indik, Z.; Yeh, H.; Ornstein-Goldstein, N.; Rosenbloom, J.C.; Abrams, J. Biol. Chem. 264, 8887-8891, 1989
 A: Title: Characterization of the complete human elastin gene. Delineation of unusual
 A: Reference number: A33705; MUID: 89255358
 A: Accession: A33705
 A: Molecule type: DNA
 A: Residues: 1-2740
 R: Fazio, M.J.; Olseop, D.R.; Kaub, E.A.; Baldwin, C.T.; Indik, Z.; Ornstein-Goldstein, J. Invest. Dermatol. 91, 458-464, 1988
 A: Title: Cloning of full-length elastin cDNAs from a human skin fibroblast recombinant
 A: Reference number: A30524; MUID: 89009960
 A: Accession: A30524
 A: Molecule type: mRNA
 A: Residues: 1-453, 483-617, 651-792 <X>
 A: Cross-references: EMBL:A58660; NID:G182061; PDB:AAA52382.1; PDB:G182062
 R: Fazio, M.J.; Olseop, D.R.; Kaub, E.A.; Baldwin, C.T.; Indik, Z.; Ornstein-Goldstein, J. Invest. Dermatol. 98, 270-277, 1988
 A: Title: Isolation and characterization of human elastin cDNAs, and age-associated va
 A: Reference number: A53891; MUID: 86156138
 A: Accession: A53891
 A: Molecule type: mRNA
 A: Residues: 164-453, 483-500, 507-617, 651-792 <X>
 A: Cross-references: EMBL:Z47482; NID:G182063; PDB:AAA53190.1; PDB:G182064
 C: Comment: the first tropoelastin refers to a soluble precursor form of the extracellu
 C: GenetObs: activity.
 A: Gene: GDB:ELN
 A: Cross-references: GDB:119107; OMIM:130160
 A: Map position: 7q11.23-7q11.23
 C: Superfamily: elastin
 C: Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
 F: 1-26/domain: signal sequence #status predicted <IG>
 F: 712-792/product: elastin #status predicted <MY>
 F: 762-767/Disulfide bonds: #status predicted

Query Match 100.0%; Score 31; DB 1; Length 792;
 Best Local Similarity 100.0%; Pred. NC. 1.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 510 PGVGA 515

RESULT 5
A:Accession: A45560
A:Title: sporozoite surface antigen SPAG-1 - Theileria annulata
C:Species: Theileria annulata
C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 20-Mar-1998
R:Hall, R.; Hunt, P.D.; Carrington, M.; Simmons, D.; Williamson, S.; Mechan, R.P.; Tait, M.O.; Biochem. Parasitol. 53, 105-112, 1992
A:Note: Mimicry of Sica protein repetitive motifs by Theileria annulata sporozoite surface antigen.
C:Accession: A45560
A:Title: Sporozoite surface antigen SPAG-1
C:Accession: A45560
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HAL>
A:Cross-references: GB:M63017; NID:g161884; PID:g161885
A:Note: sequence extracted from NCBI backbone (NCBI:111148, NCBI:111150)
C:Keywords: surface antigen

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 907;
Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 215 PGVGA 220

RESULT 6
B35962
A:Accession: B35962
A:Title: protein-tyrosine kinase (EC 2.7.1.112) I, splice form A - human
N:Alternate names: c-abl proto-oncogene homolog 2
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 04-Feb-2000
R:Kruh, G.D.; Perego, R.; Mikl, T.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 5802-5806, 1990
A:Title: The complete coding sequence of arg defines the Abelson subfamily of cytoplasmic tyrosine kinases.
C:Accession: B35962
A:Molecule type: mRNA
A:Residues: 1-1146 <KRU>
A:Cross-references: GB:N55296
A:Note: This is the complete coding sequence of arg defines the Abelson subfamily of cytoplasmic tyrosine kinases.
A:Map position: 1q24-lq25
A:Cross-references: GDB:119641; OMIM:164690
C:Superfamily: human protein-tyrosine kinase abl; protein kinase homology; SH2 homology; F78-126/Domain; SH3 homology <SH3>
F:137-227/Domain: SH2 homology <SH2>
F:250-510/Domain: protein kinase homology <KIN>
F:258-266/Region: protein kinase ATP-binding motif

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 1146;
Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 841 PGVGA 846

RESULT 7
A42459

paraesporal crystal protein cryIFal - Bacillus thuringiensis (strain aizawai)
N:Alternate names: paraesporal crystal protein cryIF
C:Species: Bacillus thuringiensis
C>Date: 10-Jul-1992 #sequence_revision 10-Jul-1992 #text_change 01-Dec-2000
C:Accession: A42459
R:Jany, C.S.; Johnson, T.B.; Garcon-Burke, J.; Bacteriol. 173, 3966-3976, 1991
A:Title: Isolation and characterization of a novel insecticidal crystal protein gene
C:Accession: A42459
A:Reference number: A42459; MUID:91286178
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <CHA>
A:Cross-references: GB:M63897; NID:g142757; PID:9142759
C:Superfamily: paraesporal crystal protein
C:Keywords: delta-endotoxin

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 1174;
Pred. No. 2.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 50 PGVGA 55

RESULT 8
B32649
A:Accession: B32649
A:Title: paraesporal crystal protein cryIa3 - Bacillus thuringiensis
C:Species: Bacillus thuringiensis
C>Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 01-Dec-2000
C:Accession: B32649
R:Lambert, B.
submitted to the EMBL Data Library, April 1993
A:Reference number: S32645
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1174 <DM>
A:Cross-references: EMBL:222512; NID:g295865; PID:CAA80235.1; PID:g295866
C:Superfamily: paraesporal crystal protein
C:Keywords: delta-endotoxin

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 1174;
Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 50 PGVGA 55

RESULT 9
A35962
A:Accession: A35962
A:Title: protein-tyrosine kinase (EC 2.7.1.112) I, splice form B - human
N:Alternate names: c-abl proto-oncogene homolog 2
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 04-Feb-2000
R:Kruh, G.D.; Perego, R.; Mikl, T.; Aaronson, S.A.
Proc. Natl. Acad. Sci. U.S.A. 87, 5802-5806, 1990
A:Title: The complete coding sequence of arg defines the Abelson subfamily of cytoplasmic tyrosine kinases.
C:Accession: A35962
A:Molecule type: mRNA
A:Residues: 1-1162 <KRU>
A:Cross-references: GB:M5296; NID:g178992; PID:AAA35553.1; PID:g178993
R:Kruh, G.D.; King, C.R.; Kraus, M.H.; Popescu, N.C.; Amsbaugh, S.C.; McBride, W.O.; Science 234, 1545-1548, 1986
A:Title: A novel human gene closely related to the abl proto-oncogene.
A:Reference number: A47577; MUID:87069952

```

A.Gene: ftsH; SCOREB:SCA10.18c

Query Match          95.8%; Score 30; DB 2; Length 456;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 133 PGIGA 138

RESULT 12
CR3530
potassium uptake protein kUp PA0917 [imported] - Pseudomonas aeruginosa (strain PA01)
C.Species: Pseudomonas aeruginosa
C.Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C.Accession: CR3530
R.R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.I.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
Lory, S.; Olson, M.V.
Nature 406, 955-964, 2000
A.Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A.Accession: CR3530
A.Length: 4,241,695 bp; MD5: A22550; XMD:20437537
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-634 <GTO>
A.Cross-references: GB:AE004526; GB:AE004091; MID:g9946619; FIDN:ANG04306.1; GSPDB:GN
A.Experimental source: strain PA01
C.Genetics:
A.Gene: kUp; PA0917

Query Match          95.8%; Score 30; DB 2; Length 634;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 223 PGIGA 228

RESULT 13
A70903
probable exinuclease abc subunit c - Mycobacterium tuberculosis (strain H37Rv)

```

CDate: 17-Jul-1998 #sequence_revision 17-Jul-1998 text_change 20-Jun-2000

R.Collier, S.T. Broesch, R. Parthill, J.J. Garnier, T.J. Churcher, C.J. Harris, D.J. Gordon
J. Connor, K.J. Davies, K.J. Devlin, K.J. Feltwell, T.J. Gentles, S.J. Hamlin, M.J. Holroyd,
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A.Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A.Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A.Accession: AF036007 A703607; MIMD:59253587
A.Length: 470963
A.Keywords: Nucleic acid sequence; Nucleotide sequence; Tuberculosis
A.Status: preliminary; nucleic acid sequence not shown; translation not shown
A.Molecule type: DNA
A.Residues: 1-645 <COL>
A.Cross-references: GS:Z801048; GS:ALJ24346; NID:g3356012; PID:MCA02168.1; EIO:g1542929
A.Experimental source: strain H37Kv
C.Name: IS6110
C.Superfamily: excinuclease ABC chain C

Query Match 96.8% Score 30; DB 2; Length 646;
Best Local Similarity 83.3%; Fred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QV 1 pggvgg 6
 |||||

Db 623 PGIGA 628

RESULT 14

S73913

polypeptide deformylase def - Mycoplasma pneumoniae (strain ATCC 29342)

N:Alternate names: hypothetical protein K04_Orf216

C:Species: Mycoplasma pneumoniae

A:Variety: ATCC 29342

C>Date: 27-Feb-1997 #sequence-revision 25-Apr-1997 #text-change 07-Dec-1999

C:Accession: S73913

R:Himmerreich, R.; Hilbert, H.; Plagens, H.; Firkel, E.; Li, B.C.; Herrmann, R.

Nucleic Acids Res. 24, 4420-4449, 1996

A>Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae

A:Accession: S73913

A:Accession: S73913

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-216 <EMBL>

A:Cross-references: EMBL:AB000057; GB:U00089; NID:g1674379; PID:AA896235.1; PID:g167438

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996

C:Gene: def

A:Genetic code: SGC3

C:Superfamily: polypeptide deformylase

Query Match 93.5%; Score 29; DB 2; Length 216;

Best Local Similarity 66.7%; Pred No. 1e+02; 0; Indels 0; Gaps 0;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6

Db 82 PGIGA 87

RESULT 15

G64211

conymethionine deformylase homolog - Mycoplasma genitalium

N:Alternate names: conymethionine deformylase

C>Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 07-Dec-1999

C:Accession: G64211

R:Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; M.; Fuhrmann, J.; Nguyen, D.; Utterback, T.R.; Saudek, D.M.; Phillips, C.A.; Merrick, J.; C.A.; Venter, J.C.

Science 270, 397-403, 1995

A>Title: The minimal gene complement of Mycoplasma genitalium.

A:Reference number: A64200; NID:96026346

A:Accession: G64211

A:Molecule type: DNA

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-226 <TIGR>

A:Cross-references: GB:U39590; GB:U43967; NID:g1045782; PID:g1045785; TIGR:M6105

A:Experimental source: strain G-37

C:Genetics:

A:Genetic code: SGC3

A:Start codon: GIG

C:Superfamily: polypeptide deformylase

Query Match 93.5%; Score 29; DB 1; Length 226;

Best Local Similarity 66.7%; Pred. No. 1e+02;

Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6

Db 92 PGIGA 97

Search completed: April 24, 2001, 16:42:01

Job time: 466 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

CM protein - protein search, using sw model

Run on: Apr-11 24, 2001, 16:42:48 ; Search time 44.88 Seconds
(without alignments)
4,580 Million cell updates/sec

Title: US-09-340-736-5

Perfect score: 31

Sequence: 1 FGVGVA 6

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Matched: 93435 seqs, 34255466 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Swissprot_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	730	1	ELF_HUMAN
2	31	100.0	747	1	ELF_BOVIN
3	31	100.0	1169	1	CLFB_BACTM
4	31	100.0	1174	1	CLFB_BACTM
5	31	100.0	1182	1	ABL2_HUMAN
6	30	96.8	546	1	UVRK_MYCTU
7	29	93.5	216	1	DEF_MYCSE
8	29	93.5	216	1	DEF_MYCSE
9	28	90.3	85	1	COXE_BOVIN
10	28	90.3	105	1	YK71_MYCTU
11	28	90.3	111	1	COXE_MOUSE
12	28	90.3	181	1	COXE_MOUSE
13	28	90.3	181	1	COXE_MOUSE
14	28	90.3	292	1	YK71_HUMAN
15	28	90.3	430	1	INDP_CANTR
16	28	90.3	431	1	GUDR_MOUSE
17	28	90.3	587	1	SVT3_MOUSE
18	28	90.3	588	1	SVT3_MOUSE
19	28	90.3	648	1	BGLR_CERAE
20	28	90.3	651	1	BGLR_HUMAN
21	28	90.3	690	1	PRO_DROME
22	28	90.3	777	1	TALA_POMV
23	28	90.3	782	1	TALA_POMV
24	28	90.3	785	1	TALA_POMV
25	28	90.3	1020	1	ATN2_HUMAN
26	28	90.3	1020	1	ATN2_HUMAN
27	28	90.3	1021	1	ATN1_HUMAN
28	28	90.3	1021	1	ATN1_HUMAN
29	28	90.3	1021	1	ATN1_HUMAN
30	28	90.3	1021	1	ATN1_HUMAN
31	28	90.3	1021	1	ATN1_HUMAN
32	28	90.3	1021	1	ATN1_HUMAN
33	28	90.3	1023	1	ATN1_HUMAN

34 27 87.1 76 1 ATPL_STROO P50014 streptomyce
35 27 87.1 102 1 COXE_CYPCA O13082 cyprinus ca
36 27 87.1 156 1 VGLL_PVKA P52512 pseudorabie
37 27 87.1 184 1 DEF2_BACTR O31410 bacillus st
38 27 87.1 194 1 HIS2_SALTY P10376 salmonella
39 27 87.1 195 1 HIS2_CAMOE Q9PM75 campylobact
40 27 87.1 195 1 HIS2_CAMOE Q9PM75 campylobact
41 27 87.1 195 1 HIS2_CAMOE Q9PM75 campylobact
42 27 87.1 195 1 HIS2_CAMOE Q9PM75 campylobact
43 27 87.1 199 1 HIS5_BREIN P44340 basophilus
44 27 87.1 203 1 HIS5_VIBCH Q9K5X0 vibrio chol
45 27 87.1 230 1 FLAH_METVO O06641 methanococc

ALIGNMENTS

RESULT 1	ELF_HUMAN	STANDARD:	PRT:	730 AA.
ID	ELF_HUMAN			
AC	P15502			
DT	01-APR-1990 (Rel. 14, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	ELASTIN PRECURSOR (TROPOELASTIN).			
GN	ELN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_Taxid=9606;			
RA	SEQUENCE FROM N.A.			
RC	TISSUE-Skin fibroblast.			
RC	MEDLINE=89094960; PubMed=3171221;			
RA	Fazio M.J., Olsen D.R., Kaul E.A., Baldwin C.T., Indik Z.,			
RA	Ornstein-Goldstein N., Yeh H., Rosenbloom J., Uitto J.;			
RT	"Cloning of full-length elastin cDNAs from a human skin fibroblast			
RT	utilizing exon-specific oligonucleotides";			
RL	J. invest. Dermatol. 91:458-464(1988).			
RA	SEQUENCE OF 603-730 FROM N.A.			
RC	TISSUE-Pituitary gland Hypophysis;			
RC	MEDLINE=96291399; PubMed=8689688;			
RA	Frangiskakis J.M., Ewart A.K., Morris C.A., Mervis C.B.,			
RA	Bertrand J., Robinson B.F., Klein B.P., Ensing G.J., Everett L.A.,			
RA	Green B.D., Proeschel C., Gutowski N.J., Noble M., Atkinson D.L.,			
RA	Odelberg S.J., Keating M.T.;			
RT	"LIM-kinase hemizygosity implicated in impaired visuospatial			
RT	constructive cognition";			
RL	Cell 86:59-69(1996).			
CC	-1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND			
CC	NUCHAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.			
CC	-1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER			
CC	BY HYDROPHOBIC INTERACTIONS AND BY COVALENTLY LINKED			
CC	-1- SUBCELLULAR LOCATION: CYTOSOL; EXTRACELLULAR MATRIX OF ELASTIC FIBERS.			
CC	-1- DISEASE: DELETED IN WILLIAMS-BUREN SYNDROME (WBS). A			
CC	DEVELOPMENTAL DISORDER CAUSED BY HAPLOINSUFFICIENCY FOR GENES IN A			
CC	2-CM REGION OF CHROMOSOME BAND 7Q11.23. HEMIZYGOUS DELETION MAY			
CC	CONTRIBUTE TO CERTAIN DEFECTS SUCH AS HYPERCALCEMIA AND GROWTH			
CC	DELAY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	certified and its use is not intended for commercial purposes. If you			
CC	or send an email to license@isb.sib.ch/			
CC	-----			
DR	EMBL; X15603; CAA33627.1; -			
DR	EMBL; M36860; AAA52382.1; -			

or send an email to license@ish-eth.ch).

```
CC      EMEL: J02717; AAA30503.1; -
CD      EMEL: K03505; AAA30505.1; -
DR      EMEL: K03506; AAA30506.1; -
DR      EMEL: J02855; AAA30776.1; -
DR      EMEL: M49652; AAA03519.1; -
DR      FIR: A26728; A26728.
DR      FIR: A26728; B26728.
DR      FIR: C26728; C26728.
DR      RSPR: N04242; NWK.
DR      KW      Connective tissue; Repeat; Signal;
KW      Alternative splicing.
FT      FT SIGNAL    1 26
FT      CHAIN       27 747
FT      MOD_RES     105 105 ELASTIN.
FT      MOD_RES     109 109 OXIDATIVE DEMINATION.
FT      MOD_RES     252 252 OXIDATIVE DEMINATION.
FT      MOD_RES     271 271 OXIDATIVE DEMINATION.
FT      MOD_RES     275 275 OXIDATIVE DEMINATION.
FT      MOD_RES     324 324 OXIDATIVE DEMINATION.
FT      MOD_RES     324 324 OXIDATIVE DEMINATION.
FT      MOD_RES     400 400 OXIDATIVE DEMINATION.
FT      MOD_RES     400 400 OXIDATIVE DEMINATION.
FT      MOD_RES     404 404 OXIDATIVE DEMINATION.
FT      MOD_RES     407 407 OXIDATIVE DEMINATION.
FT      MOD_RES     448 448 OXIDATIVE DEMINATION.
FT      MOD_RES     489 489 OXIDATIVE DEMINATION.
FT      MOD_RES     493 493 OXIDATIVE DEMINATION.
FT      MOD_RES     544 544 OXIDATIVE DEMINATION.
FT      MOD_RES     552 552 OXIDATIVE DEMINATION.
FT      MOD_RES     552 552 OXIDATIVE DEMINATION.
FT      MOD_RES     609 609 OXIDATIVE DEMINATION.
FT      MOD_RES     645 645 OXIDATIVE DEMINATION.
FT      MOD_RES     649 649 OXIDATIVE DEMINATION.
FT      MOD_RES     685 685 OXIDATIVE DEMINATION.
FT      MOD_RES     688 688 OXIDATIVE DEMINATION.
FT      VARSPLIC   226 239 MISSING (IN ELASTIN B).
FT      VARSPLIC   226 259 MISSING (IN ELASTIN C).
FT      CONFLICT   2 3 ES -> AG (IN REF. 2 AND 3);
FT      CONFLICT   12 12 E -> G (IN REF. 2 AND 3);
SQ      SEQUENCE   747 AA; 64229 MW; 633C0CEA1643D83 CRC64;

Query Match          100.0%; Score 31; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PEGVGA 6
        |||||
DB      501 PEGVGA 506

RESULT 3
CIFP_BACTM STANDARD; PRG: 1169 AA.
ID      CLIFP_BACTM Q9RC19; RC
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DE      PEPTICIDAL CRISTAL PROTEIN WITH ANTI-INFLAMMATORY AND ANTITUMOR ACTIVITY
DE      CYLIX OR CYLIX(B) OR CYLIX(S) L132 KDA CRYSTAL PROTEIN.
OS      Bacillus thuringiensis (subsp. morrisonii).
OC      Bacteria; Firmicutes; Bacillales; Clostridium group;
OC      Bacillus/Scaphylococcus group; Bacillus.
RN      NCBI_TxId:1441;
RP      SEQUENCE FROM N.A.
RP      Song F.; Zhang J.; Ding Z.; Chen T.; Li G.; Huang D.;
RP      et al. Crystal structure of the peptidase domain of
RP      Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RL      121
```

RP SEQUENCE FROM N.A.
 RC STRAIN-INA67;
 RL Masuda K., Asano S.;
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AF062350; AAF21767.1; -;
 DR EMBL; AB012288; BAA25298.1; -;
 DR HSSP; P02965; ICYI;
 DR InterPro; IPR001178; -;
 DR Pfam; PF00555; endotoxin; 1;
 KW Toxin; Sporulation;
 CC CONFLICT 38 38 V -> I (IN REF. 2);
 CC CONFLICT 39 39 G -> A (IN REF. 2);
 CC CONFLICT 1002 1002 G -> I (IN REF. 2);
 CC CONFLICT 1012 1012 E -> K (IN REF. 2);
 CC CONFLICT 1053 1053 MISSING (IN REF. 2);
 CC CONFLICT 1085 1085 G -> R (IN REF. 2);
 FT SEQUENCE 1169 AA; 132284 MW; BFCARF6FEET7C3LE CRC64;

Query Match 100.0%; Score 31; DB 1; Length 1169;
 Best Local Similarity 100.0%; Pred. NC. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PGVGA 6
 DB 50 PGVGA 55
 DB 50 PGVGA 55

RESULT 4
 AB012288 HUMAN STANDARD; PRT; 1174 AA.
 AC P42634;
 DT 01-NOV-1995 (Rel. 38, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TYROSINE-PROTEIN KINASE ABL2 (EC 2.7.1.112) (TYROSINE KINASE ARG).
 GN ABL2 OR ARG OR ABL.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CK NCBI_TaxID=9606;
 RN [1] PubMed=2198571;
 RX MEDLINE=00332670; PubMed=2198571;
 PY Kuhl G.D., Perago R., Wink T., Aronson S.A.;
 RA "The complete coding sequence of arg defines the Abelson subfamily of
 RT cytoplasmic tyrosine kinases.";
 CC Proc. Natl. Acad. Sci. U.S.A. 87:5802-5806(1990).
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A AND 1B (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC SUBDOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

RT genes encoding novel lepidopteran-active toxins.";
 RL Patent number US5188960, 23-FEB-1993.
 CC -1- FUNCTION: PROMOTES COLLOIDSMOTIC LYSIS BY BINDING TO THE MIDGUT
 CC EPITHELIAL CELLS OF MANY LEPIDOPTERAN LARVAE.
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING
 CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART
 CC OF THE SPORE COAT.
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE
 CC N-TERMINUS.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M33897; A322348.1; -;
 DR EMBL; M73254; A322347.1; -;
 DR HSSP; P02965; ICYI;
 DR InterPro; IPR001178; -;
 DR Pfam; PF00555; endotoxin; 1;
 KW Toxin; Sporulation.
 CC CONFLICT 1174 AA; 133621 MW; B51B9751D7F91C61 CRC64;

Query Match 100.0%; Score 31; DB 1; Length 1174;
 Best Local Similarity 100.0%; Pred. NC. 1.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PGVGA 6
 DB 50 PGVGA 55
 DB 50 PGVGA 55

RESULT 5
 AB012288 HUMAN STANDARD; PRT; 1182 AA.
 AC P42634;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE TYROSINE-PROTEIN KINASE ABL2 (EC 2.7.1.112) (TYROSINE KINASE ARG).
 GN ABL2 OR ARG OR ABL.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 CK NCBI_TaxID=9606;
 RN [1] PubMed=2198571;
 RX MEDLINE=00332670; PubMed=2198571;
 PY Kuhl G.D., Perago R., Wink T., Aronson S.A.;
 RA "The complete coding sequence of arg defines the Abelson subfamily of
 RT cytoplasmic tyrosine kinases.";
 CC Proc. Natl. Acad. Sci. U.S.A. 87:5802-5806(1990).
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP +
 CC PROTEIN TYROSINE PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1A AND 1B (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC SUBDOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U39691; AAC11324.1; AUT_INIT.
 DR HSSP; P27251; 1DEF.
 DR TIGR; MG106; .
 DR InterPro; IPR000181; .
 DR Pfam; PF01327; Peg_Deformylase; 1.
 KW Protein biosynthesis; Hydrolase; Zinc.
 FT METAL 134 134 ZINC (BY SIMILARITY).
 FT METAL 178 178 ZINC (BY SIMILARITY).
 FT ACT_SITE 179 179 BY SIMILARITY.
 FT METAL 182 182 ZINC (BY SIMILARITY).
 FT ACT_SITE 182 182 ZINC (BY SIMILARITY).
 SS SEQUENCE 216 AA; 25171 MW; 520226FA58BEEFC5 CRC64;
 SS
 SS SEQUENCE 216 AA; 25171 MW; 520226FA58BEEFC5 CRC64;
 SS

Query Match 93.5%; Score 29; DB 1; Length 216;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 82 PGIGIA 87

RESULT 8
 DEF_MCPN STANDARD; PRT; 216 AA.
 AC 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE POLYPEPTIDE DEFORMYLASE (EC 3.5.1.31) (PDF) (FORMYLMETHIONINE
 DE DEFORMYLASE).
 GN DEF OR MPN246 OR MP587.
 OS Mycoplasma pneumoniae.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasma pneumoniae.
 CC NCBI_TaxID=2104.
 QY (1)
 SS SEQUENCE FROM N.A.
 RX STRAIN-ATCC 29342 / M129;
 RX MEDLINE=97105885; PubMed=8948633;
 RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
 RA Herrmann R.;
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
 RT pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC -!- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
 CC NEMILY SYNTHESIZED PROTEINS (BY SIMILARITY).
 CC -!- CATALYTIC ACTIVITY: N-FORMYL-L-METHIONINE + H(2)O -> FORMATE +
 CC L-METHIONINE.
 CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AB000057; AAB96235.1; .
 DR HSSP; P27251; 1DEF.
 DR InterPro; IPR000181; .
 DR Pfam; PF01327; Peg_Deformylase; 1.

KW Protein biosynthesis; Hydrolase; Zinc.
 FT METAL 134 134 ZINC (BY SIMILARITY).
 FT METAL 178 178 ZINC (BY SIMILARITY).
 FT ACT_SITE 179 179 BY SIMILARITY.
 FT METAL 182 182 ZINC (BY SIMILARITY).
 FT ACT_SITE 182 182 ZINC (BY SIMILARITY).
 SS SEQUENCE 216 AA; 24592 MW; 361F43404B505DF CRC64;
 SS

Query Match 93.5%; Score 29; DB 1; Length 216;
 Best Local Similarity 66.7%; Pred. No. 58;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 82 PGIGIA 87

RESULT 9
 DEF_MCPN STANDARD; PRT; 85 AA.
 AC 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE CYTOCHROME C OXIDASE POLYPEPTIDE VIA-LIVER (EC 1.9.3.1) (SSG).
 GN COX6A1.
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 CC Bovidae; Bovinae; Bos.
 CC NCBI_TaxID=9913;
 QY (1)
 SS SEQUENCE FROM N.A.
 RX STRAIN-ATCC 29342 / M129;
 RX MEDLINE=92070527; PubMed=1720401;
 RA Ewert G.D., Zhang Y.Z., Capaldi R.A.;
 RT "Switching of bovine cytochrome c oxidase subunit VIa isoforms in
 RT skeletal muscle during development.";
 RL FEBS Lett. 292:79-84(1991).
 RN (2)
 RP SEQUENCE OF 1-32.
 RC TISSUE=Liver.
 RA MEDLINE=89006677; PubMed=284245;
 RA Yamamura W., Zhang Y.Z., Takamiya S., Capaldi R.A.;
 RT "Cytochrome c oxidase subunit VIa isoforms in heart and liver cytochrome c
 RT oxidase: Specific differences between heart and liver cytochrome c
 RT oxidase.";
 RL Biochemistry 27:4909-4914(1988).
 CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT.
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
 CC 4 FERROCYTOCHROME C.
 CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M38520; AAA30437.1; .
 DR HSSP; P27251; 1DEF.
 DR InterPro; IPR001349; .
 DR Pfam; PF0246; COX6A1; 1.
 DR OXfam; OF000000; COX6A1; 1.
 KW Cytochrome c oxidase subunit VIa.
 FT CONFLICT 19 19 L -> Y (IN REF. 2).
 FT CONFLICT 28 28 V -> L (IN REF. 2).
 FT CONFLICT 30 30 M -> T (IN REF. 2).

50 SEQUENCE 85 AA; 9507 MW; 9F493F6979E2E74B CRC64;

Query Match 90.3%; Score 28; DB 1; Length 85;
Best Local Similarity 83.3%; Pred. No. 47;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 24 PGVGS 29
|||||

RESULT 10
ID COX6A1
AC P43024
STANDARD; PRT; 105 AA.

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE VERY HYPOTHEICAL 11.1 KDA PROTEIN RV2274C.

GN MYCOBACTERIUM TUBERCULOSIS.

OS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

CC [UniTaxID=1773;

CC SEQUENCE FROM N.A.

CC MEDLINE=98295987; PubMed=9634230;

CC Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

CC Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,

CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

CC Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

CC Hornsby T., Jagels K., Krogh J., Mouton R., Murphy L.,

CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

CC Rutter S., Seeger K., Skelton S., Squares S., Sulton J.E.,

CC "Deciphering the biology of Mycobacterium tuberculosis from the

CC complete genome sequence.;

CC Nature 393:537-544(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL: 277163; CAS=00561.1;

DR TuberculList; RV2274C;

DR Hypothetical protein.

SEQUENCE 105 AA; 11046 MW; BB83716A9A9DE25 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 105;

Best Local Similarity 83.3%; Pred. No. 47;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6

DB 76 PGVGA 81

|||||

RESULT 11

ID COX6A1

AC P43024

Mus musculus (Mouse).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

[UniTaxID=10090;

CC SEQUENCE FROM N.A.

CC TISSUE=Liver;

CC MEDLINE=95178562; PubMed=7873616;

CC Grossman L.I., Rosenthal N.H., Akamatsu M., Erickson R.P.;

CC "Cloning, sequence analysis, and expression of a mouse cDNA encoding

CC cytochrome c oxidase subunit via liver isoform.;"

CC Blochim. Biophys. Acta 1260:361-364(1995).

CC [2]

CC SEQUENCE OF 2-111 FROM N.A.

CC STRAIN=BALE/C;

CC Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.

CC FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE

CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN

CC MITOCHONDRIAL ELECTRON TRANSPORT.

CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +

CC 4 FERROCYTOCHROME C.

CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.

CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

EMBL: L06465; AAA53066.1; ALF_INIT.

DR EMBL: U05440; AAA17836.1; ALF_INIT.

DR EMBL: P07471; KOCC.

DR SWISS: ZDPA02; P43024; MOUSE.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

DR EMBL: P07471; KOCC.

```

EX MEDLINE=94192956; PubMed=9144025;
FA Mell C., Seidman J., Kadenbach B.;
RT "Structural analysis of two different genes encoding liver- and
RT heart-type of cytochrome c oxidase submit VIA and a pseudogene
RT related to the COX1A-L cDNA.";
RL Gene 140:179-186(1994).
RN [2]
RP SEQUENCE OF 27-111 FROM N.A.
RC MEDLINE=89052650; PubMed=2461293;
RC MEDLINE=89052650; PubMed=2461293;
RA Schlarf A., Drete M., Winter M., Kadenbach B.;
RT "Characterization of two different genes (cDNA) for cytochrome c
RT oxidase subunit VIA from heart and liver of the rat.";
RL Biochim. Biophys. Acta 1015:368-372(1990).
RN [4]
RP SEQUENCE OF 27-48
RC TISSUE=Liver; and Brown adipose tissue;
RC TISSUE=Liver;
RC MEDLINE=90122894; PubMed=2153407;
RC MEDLINE=90122894; PubMed=2153407;
RA Kadenbach B., Stroth A., Becker A., Eckersorn C., Lottspeich F.;
RT "Tissue- and species-specific expression of cytochrome c oxidase
RT isozymes in vertebrates.";
RL Biochim. Biophys. Acta 1015:368-372(1990).
RN [4]
RP SEQUENCE OF 27-49
RC STRAIN=WISTAR; TISSUE=Liver;
RC MEDLINE=95245429; PubMed=7601105;
RC MEDLINE=95245429; PubMed=7601105;
RA Schlegel H., Koeck H., Hainke W., Brandt U., von Jagow G.;
RT "Cytochrome c oxidase in developing rat heart: Enzymic properties and
RT amino-terminal sequences suggest identity of the fetal heart and the
RT adult liver isoform.";
RL Eur. J. Biochem. 230:235-241(1995).
CC -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC MITOCHONDRIAL ELECTRON TRANSPORT.
CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC 4 FERROCYTOCHROME C.
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: X72757; CBA1067.1; AUT-SEQ.
DR EMBL: X12533; CBA1067.1; -.
DR EMBL: S81359; S81359.
DR F559; S81359.1; LOC.
DR InterPro: IPR001349.
DR PROSITE: PS01329; COX6A; 1.
KW Oxidoreductase; Inner membrane; Mitochondrion; Transit peptide.
FT TRANSIT 1 26 MITOCHONDRION.
FT CHAIN 27 111 CYTOCHROME C OXIDASE POLYPEPTIDE VIA-
FT CHAIN LIVER.
SQ SEQUENCE 111 AA; 12301 MW; 71EA06E23ED7546D CRC64;

Query Match 30.3%; Score 28; DB 1; Length 111;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGYA 6
DB 50 PGVGYA 55

RESULT 13
DEF2_BACSU STANDARD; PRT; 184 AA.
ID DEF2_BACSU

```

```

AC Q45495;
DT 30-MAR-2000 (Rel. 39; Created)
DE 30-MAR-2000 (Rel. 39; Last sequence update)
DE 30-MAR-2000 (Rel. 39; Last annotation update)
DE POLYPEPTIDE DEFORMYLASE 2 (EC 3.5.1.31) (PDF 2) (FORMYLMETHIONINE
DE DEFORMYLASE 2).
GN YKRB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RC "aldwell R.X., Ferrai E.;
RT "Sequence analysis of the mobA-ampS region of the Bacillus subtilis
RL Submitter (JULY-1997) to the EMBL/GenBank/DBJ databases
RL -!- FUNCTION: REMOVES THE FORMYL GROUP FROM THE N-TERMINAL MET OF
CC NEARLY SYNTHESIZED PROTEINS (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: N-FORMYL-L-METHIONINE + H(2)O = FORMATE +
CC L-METHIONINE.
CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE POLYPEPTIDE DEFORMYLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL: AF012285; AAC24930.1; -.
DR EMBL: Z99111; CBA1329.1; -.
DR Subtilist; Bg11815; YKRB.
DR InterPro: IPR000181; -.
DR Pfam: PF01327; Pep.deformylase; 1.
KW Protein biosynthesis; Hydrolyase; Zinc.
FT METAL 150 110 ZINC (BY SIMILARITY).
FT METAL 151 110 ZINC (BY SIMILARITY).
FT ACT-SITE 154 134 BY SIMILARITY.
FT METAL 157 157 ZINC (BY SIMILARITY).
SQ SEQUENCE 184 AA; 20655 MW; 8641BP193266C38 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 184;
Best Local Similarity 83.3%; Pred. No. 79;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGYA 5
DB 57 PGVGYA 52

RESULT 14
ID Y152_HUMAN
ID Y152_HUMAN STANDARD; PRT; 292 AA.
AC Q4165;
DT 01-NOV-1997 (Rel. 35; Created)
DT 01-NOV-1997 (Rel. 35; Last sequence update)
DT 01-OCT-2000 (Rel. 40; Last annotation update)
DE HYPOPHYSICAL PROTEIN KIAA0152.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RC MEDLINE=56127530; PubMed=8590280;
RC Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;

```

RT "Prediction of the coding sequences of unidentified human genes. IV.
 The analysis of cDNA clones from human cell line AG-1." [1]
 FL DNA Res. 2:167-174(1995)
 CC -1- SIMILARITY: TO C.ELEGANS F4432.4.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D63486; RA009773.1; -
 KW Hypothetical protein; Transmembrane
 FT TRANSMEM 270 290 POTENTIAL.
 FT DOMAIN 14 22 POLY-LEU.
 FT DOMAIN 231 238 POLY-GLU.
 SQ SEQUENCE 292 AA; 32234 MW; 448D673A5A18F09 CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 292;
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 29 PGVGA 34
 RESULT 15
 IDHE_CANTR STANDARD; PRT; 430 AA.
 AC 013285;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE ISOCITRATE DEHYDROGENASE (NADP), MITOCHONDRIAL PRECURSOR (EC 1.1.1.42)
 DE (NADP+)-SPECIFIC IDH (IDP).
 GN IDP1
 OS Candida tropicalis (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 CX NCBI_TaxID=5482;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PR233;
 RX MEDLINE=98004564; PubMed=9325427;
 RA Inajo T., Kawachi H., Atomi H., Sanuki S., Yamamoto S., Ueda M.,
 RA Tanaka A.;
 RA Immunogenetically distinct NADP-linked isocitrate dehydrogenase
 RA isozymes in mitochondria and peroxisomes of Candida tropicalis.";
 RA Arch Biochem Biophys 339:197-204(1997)
 CC -1- FUNCTION: ISOCITRATE DEHYDROGENASE (IDH) MAY REGULATE FLUX THROUGH THE
 CC TRICARBOXYLIC ACID CYCLE AND RESPIRATION. ITS PROBABLY CRITICAL
 CC FUNCTION IS THE PRODUCTION OF NADPH
 CC -1- CATALYTIC ACTIVITY: ISOCITRATE + NADP(+) = 2-OXOGLOUTARATE +
 CC CO(2) + NADPH.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
 CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE
 CC DEHYDROGENASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB004556; BAA22943.1; -

DR InterPro: IPR001804; -
 DR PROSITE: PS00470; IDH_IDH; 1.
 KW Oxidoreductase; NADP; Glyoxylate bypass; Tricarboxylic acid cycle;
 KW Transit peptide; Mitochondrion.
 FT TRANSIT 1 27 MITOCHONDRION (BY SIMILARITY).
 FT CHAIN 28 430 ISOCITRATE DEHYDROGENASE (NADP).
 FT ACT_SITE 120 120 BINDING TO ISOCITRATE (BY SIMILARITY).
 SQ SEQUENCE 430 AA; 48009 MW; 205A319496F0CCEB CRC64;
 Query Match 90.3%; Score 28; DB 1; Length 430;
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 Db 201 PGVGA 206

Search completed: April 24, 2001, 15:42:49
 Job time: 455 sec

RX MEDLINE-20195060; PubMed-10731132;
 RA Adams N.D., Ceiniker S.E., Holt R.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.M., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner S., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Boyle C., Baxter E.G., Holt G.E., Nelson C.R., Nikols G.L.G.,
 RA Raloff H., Abajian A., Adams J., Anderson J., Anderson J.,
 RA Bailey J., Bonfield W., Bork P., Bork P., Bork P., Bork P.,
 RA Baerson V., Bengtson P.P., Bermap B., Bhattacharya S., Bhat S.M.,
 RA Borkova D., Borchen M.R., Bouck J.P., Brockstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Paulos E., Delcher A., Deng Z., Meyers A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.C., Gary N.S., Gelbart W.M., Giassman K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Hanks M.L., Harvey D., Heiman T.J., Hernandez R.R., Houck C.,
 RA Holt J.W., Kaul S., Kaul S., Kaul S., Kaul S., Kaul S., Kaul S.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattle B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Nuzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.W.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

```

RA Reinhart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng H.-O.,
RA Zheng X.H., Zhong F.N., Zhou X., Zhu S., Zhu X., Smith H.-O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA The genome sequence of Drosophila melanogaster.
RT Science 287:2185-2195(2000).
RL EMBL: AB003625; AAF52823.1;
DR FLYBASE: FBgn0032156; CG13124.
SQ SEQUENCE 510 AA; 55856 MW; D6F2B3B58D15386A CRC64;

Query Match 100.0%; Score 31; DB 5; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 142 PGVGA 147

RESULT 5
QVPRMO PRELIMINARY; PRT: 540 AA.
AC QVPRMO; 2000 (T-EMBLrel. 13; Created)
DT 01-MAY-2000 (T-EMBLrel. 13; Last sequence update)
DT 01-JUN-2000 (T-EMBLrel. 14; Last annotation update)
DE CG10626 PROTEIN.
GN CG10626.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 10731132; PubMed: 10731132;
RA Adams M.D., Celisier S.E., Holt R.A., Evans C.A., Goodyne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bernos P.J., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Curtis K.C., Duan S., Dunn P., Butler H., Chaudhuri A., Chandra I.,
RA Chin C.C., Chong J., Chong Z., Davis A.D., Dev I., Dietz S.M.,
RA de Pablos B., Delcher A., Deng Z., Drenth J., Dvornak J.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laso P., Lei B., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
RA Linto P., Litalen J., Maitav N., Maitav N., Maitav N., Maitav N.,
RA Merkulov G., Mishina N., Moshayev V., Moshayev V., Moshayev V.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murthy D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nikon K., Nusketo D.R., Pacib J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinhart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,

```

```

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng H.-O.,
RA Zheng X.H., Zhong F.N., Zhou X., Zhu S., Zhu X., Smith H.-O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA The genome sequence of Drosophila melanogaster.
RT Science 287:2185-2195(2000).
RL EMBL: AB003625; AAF50775.1;
DR FLYBASE: FBgn0035610; CG10626.
DR INTERPRO: IPR000276;
DR INTERPRO: IPR000611;
DR INTERPRO: IPR001556;
DR PFAM: PF00001; 7tm1.1;
DR PRINTS: PR00237; GPCRHOOPS.
DR PRINTS: PR00358; BOMBESIN.
DR PRINTS: PR01012; NRPEPTIDE.
DR PROSITE: PS00237; G-PROTEIN-RECEPTOR.
SQ SEQUENCE 540 AA; 60612 MW; 2D2D5E1BD47D5 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 540;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 435 PGVGA 440

RESULT 6
QV5337 PRELIMINARY; PRT: 602 AA.
AC QV5337;
DT 01-JAN-1998 (T-EMBLrel. 05; Created)
DT 01-JAN-1998 (T-EMBLrel. 05; Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15; Last annotation update)
DE ELASTIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Li D.Y., Toland A.E., Boak B.B., Atkinson D.L., Ensing G.J.,
RA Morris C.A., Keating M.T.,
RA Hum. Mol. Genet. 0:0-0(1997).
DR EMBL: U93037; AAB5520.1;
DR EMBL: U93034; AAB5520.1; JOINED.
DR EMBL: U93035; AAB5520.1; JOINED.
DR EMBL: U93036; AAB5520.1; JOINED.
DR INTERPRO: IPR00104;
DR INTERPRO: IPR001179;
DR INTERPRO: IPR001451;
DR INTERPRO: IPR001451;
DR INTERPRO: IPR002195;
DR PRINTS: PR00308; NTIPRESEI.
DR PRINTS: PR00589; DOPAMINE4R.
DR PROSITE: PS00752; CLCHANNEL.
DR PROSITE: PS00101; HEXAPEPTIDASES; UNKNOWN.1.
DR PROSITE: PS00453; PKBP-PTASE.1; UNKNOWN.1.
FT NON-TER 1
FT NON-TER 602
SQ SEQUENCE 602 AA; 51807 MW; 53B5B9A71EF04807 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 602;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 462 PGVGA 467

```



```

RESULT 7
O15336 PRELIMINARY: PRT: 635 AA.
AC O15336
DT 01-JAN-1998 (TEMBLrel. 05, Created)
DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE ELASTIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DN [1] NCBI_TaxID=9606;
RA Li D.Y., Toland A.E., Boak B.B., Atkinson D.L., Ensing G.J.,
RA Morris C.A., Keating M.T.;
RL Hum. Mol. Genet. 0:0-0(1997).
DR EMBL: U93037; AA856521.1; -.
DR EMBL: U93034; AA856521.1; JOINED.
DR EMBL: U93035; AA856521.1; JOINED.
DR EMBL: U93036; AA856521.1; JOINED.
DR INTERPRO: IPR00104; -.
DR INTERPRO: IPR001179; -.
DR INTERPRO: IPR00145; -.
DR INTERPRO: IPR00140; -.
DR INTERPRO: IPR002185; -.
DR PRINTS: PR00308; ANTIPEEZEL.
DR PRINTS: PR00569; DOPAMINE4R.
DR PRINTS: PR00762; CLCHANNEL.
DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
DR PROSITE: PS00453; FBEP_FPIASE_1; UNKNOWN_1.
FT NON_TER 635
FT NON_TER 635
SQ SEQUENCE 635 AA: 55279 MW: 72950C36412B2A4 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 635;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 482 PGVGA 487

RESULT 8
Q28099 PRELIMINARY: PRT: 650 AA.
AC Q28099
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE ELASTIN-CREIL (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae.
DN [1] NCBI_TaxID=9913;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicila G., Yoon K., Rosenbloom J.;
RL Coll. Relat. Res. 7:235-247(1987).
RP SEQUENCE OF 17-35 FROM N.A.
RX MEDLINE=45280426; PubMed=2992576;
RA Cicila G., May M., Ornstein-Goldstein N., Indik Z., Morrow S.,
RA Yeh H.S., Rosenbloom J., Boyd C., Rosenbloom J., Yoon K.;
RL Biochemistry 24:3075-3080(1985).
RN [2]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=88028442; PubMed=3665402;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicila G., Yoon K., Rosenbloom J.;
RL Coll. Relat. Res. 7:235-247(1987).
RP "Sequence variation of bovine elastin mRNA due to alternative
splicing."
RL Coll. Relat. Res. 7:235-247(1987).
DR EMBL: M19372; AAA30501.1; -.
DR EMBL: M14422; AAA30501.1; JOINED.
DR EMBL: M19366; AAA30501.1; JOINED.
DR EMBL: M19367; AAA30501.1; JOINED.
DR EMBL: M19368; AAA30501.1; JOINED.
DR EMBL: M19369; AAA30501.1; JOINED.
DR EMBL: M19370; AAA30501.1; JOINED.
DR EMBL: M19371; AAA30501.1; JOINED.
DR EMBL: M22771; AAA30501.1; JOINED.
DR EMBL: M22772; AAA30501.1; JOINED.
DR EMBL: M22773; AAA30501.1; JOINED.
DR EMBL: M22774; AAA30501.1; JOINED.
DR EMBL: M22775; AAA30501.1; JOINED.
DR EMBL: M22988; AAA30501.1; JOINED.

```

```

RT splicing."
RL Coll. Relat. Res. 7:235-247(1987).
DR EMBL: M19372; AAA30499.1; -.
DR EMBL: M14422; AAA30499.1; JOINED.
DR EMBL: M19366; AAA30499.1; JOINED.
DR EMBL: M19368; AAA30499.1; JOINED.
DR EMBL: M19369; AAA30499.1; JOINED.
DR EMBL: M19370; AAA30499.1; JOINED.
DR EMBL: M19371; AAA30499.1; JOINED.
DR EMBL: M22771; AAA30499.1; JOINED.
DR EMBL: M22772; AAA30499.1; JOINED.
DR EMBL: M22773; AAA30499.1; JOINED.
DR EMBL: M22774; AAA30499.1; JOINED.
DR EMBL: M22775; AAA30499.1; JOINED.
DR EMBL: M22988; AAA30499.1; JOINED.
DR EMBL: M30100; AAA30499.1; JOINED.
DR INTERPRO: IPR000104; -.
DR PRINTS: PR00308; ANTIPEEZEL.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 650 AA: 55373 MW: CD21AB83E9076AD7 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 650;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
DB 412 PGVGA 417

RESULT 9
Q28096 PRELIMINARY: PRT: 666 AA.
AC Q28096
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
DE ELASTIN-CREIL (FRAGMENT).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae.
DN [1] NCBI_TaxID=9913;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicila G., Yoon K., Rosenbloom J.;
RL Coll. Relat. Res. 7:235-247(1987).
RP SEQUENCE OF 17-35 FROM N.A.
RX MEDLINE=85280426; PubMed=2992576;
RA Cicila G., May M., Ornstein-Goldstein N., Indik Z., Morrow S.,
RA Yeh H.S., Rosenbloom J., Boyd C., Rosenbloom J., Yoon K.;
RL Biochemistry 24:3075-3080(1985).
RN [2]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=88028442; PubMed=3665402;
RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
RA Rosenbloom J.C., Cicila G., Yoon K., Rosenbloom J.;
RL Coll. Relat. Res. 7:235-247(1987).
RP "Sequence variation of bovine elastin mRNA due to alternative
splicing."
RL Coll. Relat. Res. 7:235-247(1987).
DR EMBL: M19372; AAA30501.1; -.
DR EMBL: M14422; AAA30501.1; JOINED.
DR EMBL: M19366; AAA30501.1; JOINED.
DR EMBL: M19367; AAA30501.1; JOINED.
DR EMBL: M19368; AAA30501.1; JOINED.
DR EMBL: M19369; AAA30501.1; JOINED.
DR EMBL: M19370; AAA30501.1; JOINED.
DR EMBL: M19371; AAA30501.1; JOINED.
DR EMBL: M22771; AAA30501.1; JOINED.
DR EMBL: M22772; AAA30501.1; JOINED.
DR EMBL: M22773; AAA30501.1; JOINED.
DR EMBL: M22774; AAA30501.1; JOINED.
DR EMBL: M22775; AAA30501.1; JOINED.
DR EMBL: M22988; AAA30501.1; JOINED.

```

DR INTERPRO: IPR00104; -
 DR PRINTS: PR00308; ANTIPEPSEI.
 FT NON_TER 1
 SQ SEQUENCE 666 AA; 56435 MW; BC95862632BE1F71 CRC64;

Query Match 100.0%; Score 31; DB 6; Length 666;
 Best Local Similarity 100.0%; Pval N. 3.7e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 420 PGVGA 425

RESULT 10
 Q28098 Q28097 PRELIMINARY: PRT: 579 AA.
 ID Q28097
 AC Q28097
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE ELASTIN-GBEL2 (FRAGMENT).
 OS Bos taurus (Bovine).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE OF 17-35 FROM N.A.
 RA Cilia G., Yoon K., Ornstein-Goldstein N., Indik Z., Morrow S.,
 RA Cilia G., Yoon K., Ornstein-Goldstein N., Indik Z., Morrow S.,
 RA Yeh H.S., Rosenbloom J., Boyd C., Rosenbloom J., Yoon K.,
 RA "Structure of the 3' portion of the bovine elastin gene."
 RL Biochemistry 24:3075-3080(1985).
 RN [2]
 RP SEQUENCE OF 1-35 FROM N.A.
 RA MEDLINE=8028442; PubMed=3665402;
 RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Cilia G., Yoon K., Rosenbloom J.,
 RA "Splicing variation of bovine elastin mRNA due to alternative
 ET splicing."
 RL Coll. Relat. Res. 7:235-247(1987).

DR EMBL: M19372; AAA30500.1; JOINED.
 DR EMBL: M14422; AAA30500.1; JOINED.
 DR EMBL: M19366; AAA30500.1; JOINED.
 DR EMBL: M19367; AAA30500.1; JOINED.
 DR EMBL: M19368; AAA30500.1; JOINED.
 DR EMBL: M19369; AAA30500.1; JOINED.
 DR EMBL: M19370; AAA30500.1; JOINED.
 DR EMBL: M19371; AAA30500.1; JOINED.
 DR EMBL: M19372; AAA30500.1; JOINED.
 DR EMBL: M22772; AAA30500.1; JOINED.
 DR EMBL: M22773; AAA30500.1; JOINED.
 DR EMBL: M22774; AAA30500.1; JOINED.
 DR EMBL: M22988; AAA30500.1; JOINED.
 DR INTERPRO: IPR00104; -
 DR INTERPRO: IPR001459; -
 DR PRINTS: PR00308; ANTIPEPSEI.
 DR PRINTS: PR00959; MEVGALKINASE.
 FT NON_TER 1
 SQ SEQUENCE 679 AA; 57652 MW; EB3C019E3BD7618D CRC64;

Query Match 100.0%; Score 31; DB 6; Length 679;
 Best Local Similarity 100.0%; Pval N. 3.7e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 6
 DB 446 PGVGA 451

RESULT 11
 Q14235 Q14235 PRELIMINARY: PRT: 687 AA.
 ID Q14235
 AC Q14235
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE ELASTIN.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87274906; PubMed=3038460;
 RA Indik Z., Yoon K., Morrow S.D., Cilia G., Rosenbloom J.,
 RA Rosenbloom J., Ornstein-Goldstein N.,
 RA "Structure of the 3' region of the human elastin gene; great abundance
 ET of Alu repetitive sequences and few coding sequences."
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=87289668; PubMed=3039501;
 RA Indik Z., Yeh H., Ornstein-Goldstein N., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Peltonen L., Rosenbloom J.,
 RA "Alternative splicing of human elastin mRNA indicated by sequence
 ET analysis of cloned genomic and complementary DNA."
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
 DR EMBL: M17282; AAC98393.1; JOINED.
 DR EMBL: M19983; AAC98393.1; JOINED.
 DR EMBL: M17283; AAC98393.1; JOINED.
 DR EMBL: M17284; AAC98393.1; JOINED.
 DR EMBL: M17285; AAC98393.1; JOINED.
 DR EMBL: M17286; AAC98393.1; JOINED.
 DR EMBL: M17287; AAC98393.1; JOINED.
 DR EMBL: M17288; AAC98393.1; JOINED.
 DR EMBL: M17289; AAC98393.1; JOINED.
 DR EMBL: M17290; AAC98393.1; JOINED.
 DR EMBL: M17291; AAC98393.1; JOINED.
 DR EMBL: M17292; AAC98393.1; JOINED.
 DR EMBL: M17293; AAC98393.1; JOINED.
 DR EMBL: M17294; AAC98393.1; JOINED.
 DR EMBL: M17295; AAC98393.1; JOINED.
 DR EMBL: M17296; AAC98393.1; JOINED.
 DR EMBL: M17297; AAC98393.1; JOINED.
 DR EMBL: M17298; AAC98393.1; JOINED.
 DR EMBL: M17299; AAC98393.1; JOINED.
 DR EMBL: M17300; AAC98393.1; JOINED.
 DR INTERPRO: IPR001179; -
 DR INTERPRO: IPR001451; -
 DR PROSITE: PS00101; HEXAPEP_TRANSFERASES; UNKNOWN_1.
 DR PROSITE: PS00453; FBSP_PPIASE_1; UNKNOWN_1.
 SQ SEQUENCE 687 AA; 59579 MW; 9D5AC3C4D9F9E9E CRC64;

Query Match 100.0%; Score 31; DB 4; Length 687;
 Best Local Similarity 100.0%; Pval N. 3.7e-02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGVGA 5
 DB 456 PGVGA 461

RESULT 12
 Q28098 Q28098 PRELIMINARY: PRT: 707 AA.
 ID Q28098
 AC Q28098
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE ELASTIN (PPA) (GENE).
 OS Bos taurus (Bovinae).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;

RN PP SEQUENCE OF 17-35 FROM N.A.
 RX MEDLINE-8528426; PubMed-2992576;
 RA Cicilia G., Yoon K., Ornstein-Goldstein N., Indik Z., Morrow S.,
 RA Yeh H.S., Rosenbloom J., Boyd C., Rosenbloom J., Yoon K.;
 RT "Structure of the 3' portion of the bovine elastin gene.";
 RN Biochemistry 24:3075-3080(1985).
 [2]
 RP SEQUENCE OF 1-35 FROM N.A.
 RX MEDLINE-88028442; PubMed-3665402;
 RA Yeh H., Ornstein-Goldstein N., Indik Z., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Cicilia G., Yoon K., Rosenbloom J.;
 RT "Sequence variation of bovine elastin mRNA due to alternative
 splicing.";
 RN Biochem Biophys Res. 7:235-247(1987).
 DR ENBL; M19372; AAA30498.1; JOINED.
 DR ENBL; M11422; AAA30498.1; JOINED.
 DR ENBL; M19366; AAA30498.1; JOINED.
 DR ENBL; M19367; AAA30498.1; JOINED.
 DR ENBL; M19368; AAA30498.1; JOINED.
 DR ENBL; M19369; AAA30498.1; JOINED.
 DR ENBL; M19370; AAA30498.1; JOINED.
 DR ENBL; M19371; AAA30498.1; JOINED.
 DR ENBL; M22771; AAA30498.1; JOINED.
 DR ENBL; M22772; AAA30498.1; JOINED.
 DR ENBL; M22773; AAA30498.1; JOINED.
 DR ENBL; M22774; AAA30498.1; JOINED.
 DR ENBL; M22775; AAA30498.1; JOINED.
 DR ENBL; M22988; AAA30498.1; JOINED.
 DR ENBL; M23010; AAA30498.1; JOINED.
 DR INTERPRO: IPR001049; .
 DR PRINTS: PR00308; AMTIFREKZEEI.
 DR PRINTS: PR00309; NEVGLAKINASE.
 FT MONTER 1
 SQ SEQUENCE 707 AA; 60346 MW; PDPF559BAB34CE33 CRC64;
 Query Match 100.0%; Score 31; DB 6; Length 707;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 DB 446 PGVGA 451
 RESULT 13
 ID Q9XC22 PRELIMINARY; PPT: 722 AA.
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE ALPHA CHAIN (RC 1.17.4.1).
 GN NRDE.
 OS Mycoplasma gallisepticum.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 CC Mycoplasmataceae; Mycoplasma.
 CX NCBI_TaxID-2096;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A5969;
 RA Skamrov A., Goldman M., Feoktistova E., Baabekashvili R.;
 RA Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 RL "1'-5' CAPPING ENZYME IN PROTOZOANS: NECESSARY FOR DNA SYNTHESIS.
 CC "1'-5' CAPPING ENZYME IN PROTOZOANS: NECESSARY FOR DNA SYNTHESIS.
 CC THIOREDOXIN + H(2)O - RIBONUCLEOSIDE DIPHOSPHATE + REDUCED
 CC THIOREDOXIN
 CC -1- PATHWAY: FIRST REACTION IN THE DNA REPLICATION PATHWAY.
 DR ENBL; AF152114; AAD45273.1; .
 DR INTERPRO: IPR000788; .
 DR PFAM: PF00317; ribonucleo_red; 1.

DR PRINTS: PR01183; RIBORDASEMI.
 DR PROSITE: PS00089; RIBORED LARG; 1.
 KW Oxidoreductase; DNA replication.
 SQ SEQUENCE 722 AA; 81837 MW; C15B17ABF976B330 CRC64;
 Query Match 100.0%; Score 31; DB 2; Length 722;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PGVGA 6
 DB 245 PGVGA 250
 RESULT 14
 ID Q14233 PRELIMINARY; PPT: 724 AA.
 AC Q14233; Q14238;
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
 DE ELASTIN.
 GN ELN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID-9606;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87274906; PubMed-3038450;
 RA Indik Z., Yoon K., Morrow S.D., Cicilia G., Rosenbloom J.,
 RA Rosenbloom J., Ornstein-Goldstein N.;
 RT "Structure of the 3' region of the human elastin gene: great abundance
 of Alu repetitive sequences and few coding sequences.";
 RL Connect. Tissue Res. 16:197-211(1987).
 [2]
 RN PP SEQUENCE FROM N.A.
 RP MEDLINE-87294698; PubMed-3039501;
 RA Rosenbloom J., Ornstein-Goldstein N., Sheppard P., Anderson N.,
 RA Rosenbloom J.C., Yoon K., Rosenbloom J.;
 RT "Alternative splicing of human elastin mRNA indicated by sequence
 analysis of cloned genomic and complementary DNA.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:5680-5684(1987).
 [3]
 RP SEQUENCE OF 164-724 FROM N.A.
 RC TISSUE-PLACENTA;
 RX MEDLINE-88156138; PubMed-2831431;
 RA Fazio M.J., Olsen D.R., Kuivaniemi H., Chu M.L., Davidson J.M.,
 RA Rosenbloom J., Utito J.;
 RT "Isolation and characterization of human elastin cDNAs, and age-
 associated variation in elastin gene expression in cultured skin
 fibroblasts.";
 EL Lab Invest. 58:270-277(1988).
 DR ENBL; M17282; AAC98394.1; JOINED.
 DR ENBL; M16983; AAC98394.1; JOINED.
 DR ENBL; M17265; AAC98394.1; JOINED.
 DR ENBL; M17266; AAC98394.1; JOINED.
 DR ENBL; M17267; AAC98394.1; JOINED.
 DR ENBL; M17268; AAC98394.1; JOINED.
 DR ENBL; M17270; AAC98394.1; JOINED.
 DR ENBL; M17271; AAC98394.1; JOINED.
 DR ENBL; M17272; AAC98394.1; JOINED.
 DR ENBL; M17273; AAC98394.1; JOINED.
 DR ENBL; M17274; AAC98394.1; JOINED.
 DR ENBL; M17275; AAC98394.1; JOINED.
 DR ENBL; M17276; AAC98394.1; JOINED.
 DR ENBL; M17277; AAC98394.1; JOINED.
 DR ENBL; M17278; AAC98394.1; JOINED.
 DR ENBL; M17279; AAC98394.1; JOINED.
 DR ENBL; M17280; AAC98394.1; JOINED.
 DR ENBL; M17281; AAC98394.1; JOINED.
 DR ENBL; M24782; AAA53190.1; .
 DR INTERPRO: IPR001179; .

DR INTERPRO: IPR001451; --
 DR PROSITE: PS00101; HEXAPEP TRANSFERASES; UNKNOWN 1.
 DR PROSITE: PS00453; FBEP_PPASE_1; UNKNOWN 1.
 SQ SEQUENCE 734 AA: 62664 MW: 61155686228ED3D CRC64;

Query Match 100.0%; Score 31; DB 4; Length 724;
 Best Local Similarity 100.0%; Pred. No. 3.9e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 PGVGA 6
 Db 475 PGVGA 480

RESULT 15

QVWNI PRELIMINARY; PRT: 747 AA.

QVWNI: 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequences update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG12527 PROTEIN.
 GN CG12527
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRK1567; PubMed=10711132;
 RA Adams W.D., Celisner S.G., Holt K.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang X., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe X., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beksley E.M.,
 RA Beeson K.I., Bemis P.V., Bertman B.P., Bhandari D., Bolshakov S.,
 RA Borker N., Borker A.A., Borker R., Brockstein P., Bottler J.,
 RA Burtis K.C., Busch A.A., Butler R.C., Cavagnat L.F., Davies P.,
 RA Cherry J.M., Cawley S., Dahke C., Davapont L.E., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Davis A.D., Dew I., Dierz S.G.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dankov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Rowland T.J., Wei M.-B., Ibegwam C.,
 RA Jaisail M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko F., Leal T., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Lisso P., Little B., McCluskey M.P., McPherson J.,
 RA Markov G., Milashin N., Morley C., Murray D.K., Nadeau A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muray D.K., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner E., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Zhu J., Zhu R., Zhu S., Zhu M., Zhu M., Zhu G., Zhao Q., Zheng D.,
 RA Zhu X.B., Zhao F., Zhu X., Zhu X., Zhu X., Zhu X., Zhu X.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RA The genome sequence of Drosophila melanogaster.
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003510; AF048907.1;
 DR FLYBASE: FBgn0030377; CG12527.

SQ SEQUENCE 747 AA: 76047 MW: 3008028A92C4EEF1 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 747;
 Best Local Similarity 100.0%; Pred. No. 4.1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 PGVGA 6
 Db 680 PGVGA 685

Search completed: April 24, 2001, 16:40:35
 Job time: 425 sec



CC linking component. The copolymer is useful in prosthetic systems, for
 CC retaining elasticity. It is a functionalized polymer, as to
 CC provide reactive groups which become covalently cross-linked by
 CC tissue enzymes to newly synthesized connective tissue protein.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 7; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VFGVG 5
 Db 1 VPGVG 5

RESULT 2
 R29149
 ID R29149 standard; peptide: 5 AA.

XX R29149;

06-MAY-1993 (first entry)

XX Pentapeptide repeating unit of bioelastic polymer.

DE Superabsorbent; bioelastic; diaper; hygienic articles; wound;
 XX dressing; implant; inverse temperature transition; tissue; napkin;
 KW carbox liner; toilet paper; towellette; cleaning wipe; bandage;
 KW medical sponge; swab; printing ink; contact lens.

XX Synthetic.

OS WO9218079-A.

PN 29-OCT-1992.

XX 10-MAR-1992; 92WO-US01959.

XX 19-APR-1991; 91US-0688185.

XX (BIOS-) BIOELASTICS RES LTD.

XX Urry DW;

WPI; 1992-381725/46.

XX Super-absorbent material incorporating polymer undergoing inverse
 PT temp. transition - esp. bio-elastic polypeptide(s) for
 CC controllably absorbing body fluids

PS Claim 27/28; Page 34; 42pp; English.

XX The invention relates to an appliance suitable for contacting body
 CC surfaces and for absorbing aqueous liquids including body fluids.
 CC The appliance includes a polymeric material which undergoes an
 CC inverse temperature transition, pref. such that it is in a
 CC contracted state at a higher temperature and in a swollen state when
 CC at a lower temperature. The appliance, e.g. a diaper, may be
 CC selected to be in a contracted state prior to use and while
 CC contacting the body surface, and to be in a swollen state after
 CC absorbing a body fluid and being at a new location having a lower
 CC temperature. The invention also includes a method of using a
 CC polymeric material in a bioelastic polymer containing elastomeric
 CC tetrapeptide, pentapeptide or nonapeptide repeat units, the polymer
 CC having a series of beta turns separated by dynamic bridging segments
 CC suspended between the beta turns. In particular, the polymer comprises a
 CC segment of formula poly[(VFGXG)(VFGVG)Y], in which x and y are mole
 CC fractions such that x+y = 1, and x is a hydrophobic amino acid residue;
 CC or of formula poly[(VFGXG)(VFGVG)(VFGZG)Z], in which x+y+z = 1, x = as
 CC above, and z is an amino acid residue having a side chain capable of
 CC undergoing reversible protonation in an aqueous environment.

CC A specific polymer is poly(VFGVG).

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VFGVG 5
 Db 1 VPGVG 5

RESULT 3
 R29145
 ID R29145 standard; peptide: 5 AA.

XX R29145;

XX 23-MAR-1993 (first entry)

XX Bioelastic pentapeptide polymer fragment 2.

XX Polymer fragment; inverse temperature transition; exerted pressure;
 KW mechanical work; chemical change; hydrophobicity.

OS Synthetic.

XX WO9219183-A.

XX 12-NOV-1992.

XX 03-APR-1992; 92WO-US02691.

XX 22-APR-1991; 91US-0688324.

XX (UNAB-) UNAB RES FOUND.

XX Urry DW;

WPI; 1992-398459/48.

XX Pressure expanding polymers giving reversible mechanical changes
 PT - has inverse temp. transition in liq. water range and pref.
 CC being bio-elastic polypeptide(s) contg. hydrophobic gps.

XX Claims 13 and 14; Page 36; 51pp; English.

XX The sequences given in R29144-46 are polymer fragments that are used
 CC in polymeric materials which have an inverse temperature transition in the
 CC water range and which are used in a contracted state and a swollen state.
 CC Fragments 1 and 2 are used such that mole fractions equal 1 in one
 CC polymeric material and fragments 1, 2 and 3 are used in a further
 CC polymeric material such that mole fractions equal 1. The materials
 CC can be used in a variety of applications to produce mechanical work
 CC and/or cause chemical changes in a sealed environment by variation of
 CC the pressure on the material. The degree of mechanical or chemical
 CC change can be controlled by selection of the number, hydrophobicity
 CC and size of the hydrophobic groups and the presence or absence of
 CC reactive functional groups in the polymer.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 13; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VFGVG 5
 Db 1 VPGVG 5

```

RESULT 4
R5228
ID R5228 standard; peptide: 5 AA.
XX
AC R5228;
XX
DT 12-OCT-1995 (first entry)
XX
DE Fibrinogen/beta-casein transglutaminase cross-linkable copolymer.
XX
KW transglutaminase cross-linkable copolymer; fibrinogen; beta-casein;
KW biocompatible; bioadhesive polymers; wound repair; prosthetics;
KW bone and soft tissue matrices; controlled drug release carriers.
XX
OS Synthetic.
XX
WO9505396-A.
XX
23-FEB-1995.
XX
05-AUG-1994; 94WO-US08754.
XX
13-AUG-1993; 93US-0106509.
XX
(2YMO) ZYMOGENETICS INC.
XX
Busby SJ, Labroc VM;
XX
WPI, 1995-09872/13.
XX
New peptide(s) based on fibrinogen and beta-casein.
XX
cross-linkable by transglutaminase, used for preparing
XX
biocompatible, bioadhesive polymers
XX
Claim 16; Page 48; 59pp; English.
XX
A biocompatible, bioadhesive, transglutaminase cross-linkable
XX
copolymer comprises a first polypeptide monomer from 13-120 amino
XX
acids containing a segment of the formula SI-Y-S2 cross-linkable
XX
by a transglutaminase, where SI = R5218, S2 = R5219 and Y = a
XX
second polypeptide, where R5218 is defined as R5220 and
XX
second polypeptide is defined as R5221, R5222, R5223, R5224,
XX
first monomer can also be R5227. The copolymer can be used to
XX
produce tissue adhesives, wound repair formulations, rigid to
XX
prosthetics, matrices for the replacement of bone and soft tissue
XX
structures and carriers for controlled drug release compans.
XX
Sequence 5 AA:
SQ
Query Match 100.0%; Score 27; DB 16; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
DB 1 VPGVG 5
RESULT 6
R80307
ID R80307 standard; peptide: 5 AA.
XX
AC R80307;
XX
17-APR-1996 (first entry)
XX
Elastin cross-linking substrate consensus sequence.
XX
pendent group; repeating unit; enzyme recognition site; sealant; elastin;
XX
enzymatic cross-linking; biocompatible material; structural integrity;
XX
medical adhesive; wound closure; tissue repair.
XX
OS Synthetic.
XX
WO9523611-A1.
XX
08-SEP-1995.
XX
03-MAR-1995; 95WO-US02728.
XX
03-MAR-1994; 94US-0205518.
XX
(PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX

```


XX Cappelletto J;
 XX WPI; 1995-32043/41.
 XX
 XX Protein polymers comprising repeating units and sequences - capable
 PT of enzyme-catalysed covalent bond formation useful as a
 PT biocompatible material for wound closure and tissue repair
 XX
 XX Disclosure; Page 12: 138pp; English.
 XX
 XX The sequence of the cross-linking reactive motif from elastin. The motif
 CC can be used in a novel polymer comprising two spaced enzyme recognition
 CC site and may contain repetitive units of 3-8 amino acids with at least
 CC two pendant groups, the polymers being cross-linkable upon exposure
 CC to a biological system where in situ formation of a biocompatible
 CC material with structural integrity is required e.g. as medical adhesives
 CC and sealants or for wound closure or tissue repair.
 XX
 XX Sequence 5 AA:
 SQ
 Query Match 100.0%; Score 27; DB 16; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VPGVG 5
 Db 1 VPGVG 5
 RESULT 7
 W22714
 ID W22714 standard; peptide: 5 AA.
 AC W22714;
 XX 26-FEB-1998 (first entry)
 XX Bioelastomeric repeating unit 1 responsive to electrical energy.
 DE Bioelastomeric repeating unit; polymer; inverse temperature transition;
 KW electrical energy change; polarity; hydrophobicity; mechanical work;
 KW surgical suture; heat shrinkable membrane; desalination.
 XX Synthetic.
 OS
 XX W03723729-AL.
 XX 03-JUL-1997.
 XX 07-JUN-1996; 96W0-US09776.
 XX 07-JUN-1995; 95US-0487594.
 XX (URRX/) URRX D W.
 XX Urry DW;
 XX WPI; 1997-363360/33.
 XX Bioelastic polymer responsive to electrical energy - comprising
 PT beta turn and residue(s) with side chain that changes polarity or
 PT hydrophobicity in response to electrical energy change, useful for
 PT mechanical work or light stimulated contraction
 XX
 XX Claim 13; Page 52; 60pp; English.
 XX W22711-16 represent bioelastomeric repeating units that are used to
 CC create a polymer responsive to electrical energy. This bioelastic
 CC polymer comprises a bioelastomeric polypeptide having an inverse
 CC temperature transition, and a repeating bioelastomeric unit containing at

CC least 1 beta-turn and residues with a side chain. The side chain of these
 CC residues responds to an electrical energy change by altering its polarity
 CC hydrophobicity. The side chain is present in an amount sufficient to
 CC confer sufficient hydrophobicity to the polymer upon transition of the polymer upon
 CC the change in exposure to electrical energy. The polymer may also contain
 CC a second amino acid with a side chain capable of undergoing a change in
 CC an aqueous environment. The polymer (W34881-83) consists of the formulas
 CC described in W22716 and W22711 and below:
 CC poly{fx(VPGXG),fv(VPGVG)} where
 CC fx and fv are mole fractions with fx + fv = 1
 CC x represents an amino acid residue having an electrically responsive
 CC side chain.
 CC The expansion or contraction of the polymer upon exposure to a change
 CC in electrical energy can be used to produce mechanical work. When the
 CC polymer is used in a suture, the change in electrical energy and the changes
 CC in released environment. The polymer can be used in surgical sutures,
 CC especially for microsurgery, heat shrinkable membranes, controlled
 CC herbicide or pesticide release or for desalination.
 XX
 XX Sequence 5 AA:
 SQ
 Query Match 100.0%; Score 27; DB 18; Length 5;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 VPGVG 5
 Db 1 VPGVG 5
 RESULT 8
 W26332
 ID W26332 standard; Peptide: 5 AA.
 AC W26332;
 XX 19-NOV-1997 (first entry)
 XX Elastin repeat unit.
 DE Elastin; protein polymer.
 XX Synthetic.
 OS
 XX US5641648-A.
 XX 24-JUN-1997.
 XX 04-NOV-1986; 86US-0927258.
 XX 29-DEC-1993; 93US-0175155.
 XX 04-NOV-1986; 86US-0927258.
 XX 29-OCT-1987; 87US-0114618.
 XX 09-NOV-1988; 88US-0259429.
 XX 06-NOV-1990; 90US-0609716.
 XX 22-APR-1993; 93US-0053049.
 XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 XX Cappelletto J, Ferrari FA, Richardson C;
 XX WPI; 1997-340943/31.
 XX Preparation of synthetic DNA encoding a protein comprising short
 PT repeats - by synthesising oligomers, annealing and oligomerising
 PT these, particularly to produce proteins that mimic silk, collagen
 PT etc
 XX Claim 8; Column 11; 90pp; English.
 XX This peptide represents a repeat unit found in elastin. Methods
 CC are claimed for preparing protein polymers that contain repeating

CC units based on those found in elastin, fibroin, keratin, etc. (see
 CC also W26329, W26334 and W26336). The method involves: synthesizing
 CC pairs of single-stranded oligomers, each member of the pair
 CC overlapping except at the protruding ends; hybridising each pair to
 CC double-stranded segments; combining the double-stranded segments, or
 CC their cloned copies, in a cloning vector to form a monomer; excising
 CC the monomer from the vector; and oligomerising the monomer to
 CC produce a multimer. The multimer is then used to form a structural
 CC propeptide, including crystalline elastin, tropoelastin, or
 CC propeptides such as those that mimic (and can substitute for) silk,
 CC elastin, collagen, keratin etc. The properties of the protein can
 CC be controlled by varying the type of units in the monomer, the
 CC number of units per multimer, the spacing between them and the
 CC number of multimer repeats.

XX Sequence 5 AA;

XX Query Match 100.0%; Score 27; DB 18; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 3.2e+05;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

DB 1 VPGVG 5

XX RESULT 9

XX ID W18261 standard; peptide: 5 AA.

XX AC W18261;

XX DT 29-AUG-1997 (first entry)

XX DE Transglutaminase cross-linkable polypeptide elastomeric peptide.

XX KW Elastomeric; homopolymer; copolymer; tissue sealant; skin graft;

XX burn.

XX Synthetic.

XX PN W09640780-A1.

XX DT 19-DEC-1996.

XX PR 31-MAY-1996; 96NO-US08269.

XX PA (ZYMO) ZYMOGENETICS INC.

XX PI Busby SJ, Labroc WM;

XX DT 1997-052237/05.

XX PT Transglutaminase cross-linkable peptide(s) - used in the mfr. of
 PT biocompatible, bioadhesive tissue sealant and wound healing
 PT preparations.

XX PS Claim 6; Page 48; 55pp; English.

XX CC Novel polypeptides optionally having one or both of the amino-terminus
 CC and/or carboxy terminus flanked by an elastomeric peptide, which is cross-
 CC linkable by transglutaminase, are suggested for use in tissue sealants,
 CC SI-Y-92. The present sequence represents a specifically claimed
 CC example of an elastomeric peptide. The homo- and copolymers produced
 CC are useful in tissue sealant and wound healing formulations. Tissue
 CC sealants are useful in skin grafting for burn victims and for sealing
 CC surgical and other wounds.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

DB 1 VPGVG 5

XX RESULT 10

XX ID W12301 standard; peptide: 5 AA.

XX AC W12301;

XX DT 22-APR-1997 (first entry)

XX DE Monomeric unit for elastic protein-base polymer for chewing gum.

XX KW Chewing gum; water-insoluble; plasticiser; elastomer; flavour; polymer;
 XX sweetener; texturiser; crosslinkage; gamma-irradiation; drug; nutrient;
 XX environmentally friendly; drug; anti-inflammatory agent; vitamin.

XX OS Synthetic.

XX PN US5580390-A.

XX DT 03-DEC-1996.

XX PR 27-DEC-1993; 93US-0174185.

XX PR 27-DEC-1993; 93US-0174185.

XX PA (WRIL) WRIGLEY JR CO WM.

XX PI Hartman SE;

XX DT 1997-033531/03.

XX PT Environmentally friendly chewing gum - contg. water insoluble

PT elastic polypeptide having a penta-peptide repeat

XX PS Claim 1; Column 7; 6pp; English.

XX CC A novel chewing gum composition comprises a water insoluble polymer
 CC comprising this penta-peptide as a base. The insoluble portion can
 CC comprise 1-99% of the gum, together with a plasticiser, an elastomer,
 CC a flavour, a sweetener and a texturiser. The peptide is preferably
 CC crosslinked by gamma-irradiation to form the polymer. The new chewing
 CC gum is environmentally friendly as it can be swallowed after chewing or
 CC easily removed from surfaces. The peptide can also have drugs or other
 CC nutrients e.g. anti-inflammatory agents, vitamins, etc., attached to the
 CC peptide, for therapeutic purposes.

XX Sequence 5 AA;

Query Match 100.0%; Score 27; DB 18; Length 5;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

DB 1 VPGVG 5

XX RESULT 11.

XX ID W49701 standard; peptide: 5 AA.

XX AC W49701;

XX XX

Query Match 100.0%; Score 27; DB 20; Length 5;
Best Local Similarity 100.0%; Pred. No. 3.2e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 vpgvg 5
|||||
DB 1 vpgvg 5

Search completed: April 24, 2001, 16:38:19
Job time: 419 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:36:24 ; Search time 62.39 Seconds
(without alignments)
1.540 Million cell updates/sec

Title: US-09-340-736-6

Percent score: 27

Sequence: 1 VPGVG 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_5/prodata/2/aaa/5A_COMB.pep.*
- 2: /cgn2_5/prodata/2/aaa/5B_COMB.pep.*
- 3: /cgn2_5/prodata/2/aaa/6A_COMB.pep.*
- 4: /cgn2_5/prodata/2/aaa/6B_COMB.pep.*
- 5: /cgn2_5/prodata/2/aaa/RTGS_COMB.pep.*
- 6: /cgn2_5/prodata/2/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	27	100.0	5	1 US-08-106-509-5	Sequence 5, Appl.
2	27	100.0	5	1 US-07-609-716-2	Sequence 2, Appl.
3	27	100.0	5	1 US-08-174-185-1	Sequence 1, Appl.
4	27	100.0	5	1 US-08-212-237-2	Sequence 2, Appl.
5	27	100.0	5	1 US-08-175-153-4	Sequence 9, Appl.
6	27	100.0	5	1 US-08-468-543-15	Sequence 15, Appl.
7	27	100.0	5	1 US-08-397-2098-4	Sequence 14, Appl.
8	27	100.0	5	1 US-08-397-2098-4	Sequence 14, Appl.
9	27	100.0	5	1 US-08-397-2098-4	Sequence 14, Appl.
10	27	100.0	5	1 US-08-397-2098-4	Sequence 14, Appl.
11	27	100.0	5	2 US-08-707-237A-10	Sequence 10, Appl.
12	27	100.0	5	2 US-08-483-236-15	Sequence 15, Appl.
13	27	100.0	5	2 US-08-469-692-15	Sequence 15, Appl.
14	27	100.0	5	2 US-08-911-364-6	Sequence 6, Appl.
15	27	100.0	5	2 US-08-735-692-20	Sequence 20, Appl.
16	27	100.0	5	2 US-08-398-046-15	Sequence 15, Appl.
17	27	100.0	5	3 US-08-542-051-1	Sequence 1, Appl.
18	27	100.0	5	3 US-08-482-0859-24	Sequence 24, Appl.
19	27	100.0	5	3 US-08-482-0859-24	Sequence 43, Appl.
20	27	100.0	5	4 US-08-792-186-27	Sequence 27, Appl.
21	27	100.0	5	4 US-08-792-186-27	Sequence 27, Appl.
22	27	100.0	5	4 US-08-792-186-27	Sequence 27, Appl.
23	27	100.0	5	5 PCN-US95-02772-2	Sequence 2, Appl.
24	27	100.0	5	5 PCN-US95-02772-2	Sequence 2, Appl.
25	27	100.0	5	5 5250516-1	Patent No. 5250516
26	27	100.0	5	6 5250516-17	Patent No. 5250516
27	27	100.0	6	1 US-08-468-543-12	Sequence 12, Appl.
28	27	100.0	6	2 US-08-469-692-12	Sequence 12, Appl.

Sequence 12, Appl.
Patent No. 5250516
Sequence 17, Appl.
Sequence 17, Appl.
Sequence 17, Appl.
Sequence 35, Appl.
Sequence 35, Appl.
Sequence 35, Appl.
Sequence 35, Appl.
Sequence 35, Appl.
Sequence 35, Appl.
Sequence 14, Appl.
Sequence 14, Appl.
Sequence 14, Appl.
Sequence 16, Appl.
Sequence 2, Appl.
Sequence 14, Appl.
Sequence 14, Appl.
Sequence 16, Appl.

ALIGNMENTS

RESULT 1
US-08-106-509-5

Sequence 5, Application US/08106509

Patent No. 5428014

GENERAL INFORMATION:

APPLICANT: Labroc, Virender

APPLICANT: Busby, Sharon J.

TITLE OF INVENTION: TRANSLOCINASE CROSS-LINKABLE

TITLE OF INVENTION: POLYPEPTIDES AND METHODS RELATING THERETO

NUMBER OF INVENTORS: 14

NUMBER OF INVENTORS: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Zymogenetics, Inc.

STREET: 4225 Roosevelt Way, N.E.

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/106-509

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Parker, Gary E

REGISTRATION NUMBER: 31-684

REFERENCE/DOCID: 09

TELEPHONE: 206-547-6050 ext 322

TELEPHONE: 206-548-2329

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 5 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

US-08-106-509-5

Query Match 100.0%; Score 27; DB 1; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.4e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

DB 1 VPGVG 5

RESULT 2
US-08-609-716-2
; Sequence 1, Application US/07609716
; Patent No. 5514581
; GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.
TITLE OF INVENTION: Functional Recombinantly Prepared
TITLE OF INVENTION: Synthetic Protein Polymer
NUMBER OF SEQUENCES: 118
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/609,716
FILING DATE: 06-NOV-1990
COUNTRY: US

ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: A-55186-3/BIR
REFERENCE/DOCKET NUMBER: 20015
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-609-716-2

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 1 VPGVG 5
Db 1 VPGVG 5

RESULT 3
US-08-174-185-1
; Sequence 1, Application US/08174185
; Patent No. 5580590
; GENERAL INFORMATION:

APPLICANT: Hartman, Scott E
TITLE OF INVENTION: Environmentally Friendly
TITLE OF INVENTION: Chewing Gum Compositions Containing Elastic Protein-Based
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hill Streadman & Simpson
STREET: 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4 Mb

MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/174,185
FILING DATE: 17-DEC-1991
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acid residues
TYPE: amino acid
STRANDEDNESS: N/A
TOPOLOGY: N/A
MOLECULE TYPE: peptide
US-08-174-185-1

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 1 VPGVG 5
Db 1 VPGVG 5

RESULT 4
US-08-212-237-2
; Sequence 1, Application US/08212237
; Patent No. 558847/BIR
; GENERAL INFORMATION:
APPLICANT: Cappello, Joseph
TITLE OF INVENTION: Synthetic Proteins As Implants
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: U.S.A.
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/212,237
FILING DATE: 11-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: A-55847/BIR
REFERENCE/DOCKET NUMBER: A-55847/BIR
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-212-237-2

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

Qy 1 VPGVG 5
Db 1 VPGVG 5

Db 1 VPGVG 5

RESULT 5

US-08-175-155-4
Sequence 4, Application US/08175155
Patent No. 5641648
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
APPLICANT: Crissman, John W.
APPLICANT: Dorman, Mary A.
TITLE OF INVENTION: Methods for Preparing Synthetic
TITLE OF INVENTION: Repetitive DNA
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/175,155
FILING DATE: 29-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I.
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-5/3IR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-175-155-4

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

Db 1 VPGVG 5

RESULT 6

US-08-175-155-9
Sequence 9, Application US/08175155
Patent No. 5641648
GENERAL INFORMATION:
APPLICANT: Ferrari, Franco A.
APPLICANT: Cappello, Joseph
APPLICANT: Crissman, John W.
APPLICANT: Dorman, Mary A.
TITLE OF INVENTION: Methods for Preparing Synthetic
TITLE OF INVENTION: Repetitive DNA
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
ZIP: 94111

STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/175,155
FILING DATE: 29-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I.
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: A-55186-5/3IR
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-175-155-9

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5

Db 1 VPGVG 5

RESULT 7

US-08-468-543-15
Sequence 15, Application US/08468543
Patent No. 5726153
GENERAL INFORMATION:
APPLICANT: Lees, Robert S. et al.
TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR ARTERIAL IMAGING
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/468,543
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/398,046
FILING DATE: 02-MAR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/201,057
FILING DATE: 24-FEB-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/048,569
FILING DATE: 16-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/694,529

1 FILING DATE: 02-MAY-1991
2 PRIOR APPLICATION DATA:
3 APPLICATION NUMBER: US 07/516,215
4 FILING DATE: 04-MAY-1990
5 ATTORNEY/AGENT INFORMATION:
6 NAME: Riccardi, Richard P.
7 REGISTRATION NUMBER: 21,861
8 REFERENCE/DOCKET NUMBER: A-35186-7/RPT/ATK
9 PRIOR APPLICATION DATA:
10 FILING DATE: 03-MAY-1990
11 APPLICATION NUMBER: US 07/189,130
12 FILING DATE: 02-MAY-1988
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Clark, Paul T.
15 REGISTRATION NUMBER: 30,152
16 REFERENCE/DOCKET NUMBER: 04547/002003
17 FILING DATE: 02-MAY-1988
18 TELEPHONE: 617/542-5070
19 TELEFAX: 617/542-8906
20 TELEX: 200154
21 INFORMATION FOR SEQ ID NO: 13:
22 LENGTH: 5
23 SEQUENCE CHARACTERISTICS:
24 TYPE: amino acid
25 STRANDEDNESS: linear
26 TOPOLOGY: linear
27 US-08-466-543-15

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 1 VPGVG 5

RESULT 8
US-08-477-5098-24
1 Sequence 24, Application US/084775098
2 Patent No. 5770657
3 GENERAL INFORMATION:
4 APPLICANT: Ferrari, Franco A
5 APPLICANT: Cappello, Joseph
6 APPLICANT: Crissman, John W
7 APPLICANT: Dorman, Mary A
8 TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive
9 UNITS OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same
10 NUMBER OF SEQUENCES: 112
11 CORRESPONDENCE ADDRESS:
12 ADDRESS: Fleish, Hohnbach, Test, Albrighton & Herbert
13 STREET: Four Embarcadero Center, Suite 3400
14 CITY: San Francisco
15 STATE: California
16 COUNTRY: US
17 ZIP: 94111
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: Patent in Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA: US/08/477,509B
24 APPLICATION NUMBER: 438
25 FILING DATE: 29-DEC-1993
26 CLASSIFICATION: 438
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: US 08/175,155
29 FILING DATE: 29-DEC-1993
30 PRIOR APPLICATION DATA:
31 APPLICATION NUMBER: US 08/053,049
32 FILING DATE: 22-APR-1993
33 PRIOR APPLICATION DATA:
34 APPLICATION NUMBER: US 07/114,618
35 FILING DATE: 29-OCT-1987

1 PRIOR APPLICATION DATA:
2 APPLICATION NUMBER: US 06/927,258
3 FILING DATE: 04-NOV-1986
4 ATTORNEY/AGENT INFORMATION:
5 NAME: Riccardi, Richard P.
6 REGISTRATION NUMBER: 21,861
7 REFERENCE/DOCKET NUMBER: A-35186-7/RPT/ATK
8 TELECOMMUNICATION INFORMATION:
9 TELEPHONE: 415-781-1989
10 TELEFAX: 415-398-3249
11 INFORMATION FOR SEQ ID NO: 24:
12 SEQUENCE CHARACTERISTICS:
13 LENGTH: 5 amino acids
14 TYPE: amino acid
15 STRANDEDNESS: single
16 TOPOLOGY: linear
17 MOLECULE TYPE: Peptide
18 US-08-477-5098-24

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 1 VPGVG 5

RESULT 9
US-08-397-633A-1
1 Sequence 1, Application US/08397633A
2 Patent No. 5773577
3 GENERAL INFORMATION:
4 APPLICANT: Cappello, Joseph
5 TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
6 OF ENZYMATIC CROSS-LINKING
7 NUMBER OF SEQUENCES: 105
8 CORRESPONDENCE ADDRESS:
9 ADDRESS: FLEISH, HOHNACH, TEST, ALBRIGHTON & HERBERT
10 STREET: 4 Embarcadero Center, Suite 3400
11 CITY: San Francisco
12 STATE: California
13 COUNTRY: USA
14 ZIP: 94111-4187
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: Patent in Release #1.0, Version #1.25
20 CURRENT APPLICATION DATA:
21 APPLICATION NUMBER: US/08/397,633A
22 FILING DATE:
23 CLASSIFICATION: 530
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Rowland, Bertram I
26 REGISTRATION NUMBER: 20,015
27 REFERENCE/DOCKET NUMBER: A-36848-1/BIR PROP-011-1
28 TELECOMMUNICATION INFORMATION:
29 TELEPHONE: (415) 781-1989
30 TELEFAX: (415) 398-3249
31 TELEX: 910 277299
32 INFORMATION FOR SEQ ID NO: 1:
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 5 amino acids
35 TYPE: amino acid
36 STRANDEDNESS: single
37 TOPOLOGY: linear
38 MOLECULE TYPE: Peptide
39 US-08-397-633A-1

Query Match 100.0%; Score 27; DB 1; Length 5;

Best Local Similarity 100.0%; Prod. No. 1.4e+05; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 VPGVG 5
DB 1 VPGVG 5

RESULT 10
US-08-397-633A-8
; Sequence 10, Application US/08397633A
; Patent No. 5773577
; GENERAL INFORMATION:
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: PRODUCTS COMPRISING SUBSTRATESCAPABLE
; OF ENZYMIC CROS-LINKING
; NUMBER OF SEQUENCES: 105
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397.633A
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertam I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-58848-1/BIR PROP-011-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-397-633A-8

Query Match 100.0%; Score 27; DB 1; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 1 VPGVG 5

RESULT 11
US-08-707-237A-10
; Sequence 10, Application US/08707237A
; Patent No. 5939385
; GENERAL INFORMATION:
; APPLICANT: Ferris, Franco A.
; APPLICANT: Crissman, John W.
; APPLICANT: Dorman, Mary A.
; TITLE OF INVENTION: METHODS FOR PREPARING SYNTHETIC
; REPEATITIVE DNA
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hobach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/707.237A
; FILING DATE: 03-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION NUMBER: US 08/175,155
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/053,049
; FILING DATE: 22-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; ATTORNEY/AGENT INFORMATION:
; NAME: Tescatlin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-10/WHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-707-237A-10

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Prod. No. 1.4e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 1 VPGVG 5

RESULT 12
US-08-483-236-15
; Sequence 15, Application US/08483236
; Patent No. 5939385
; GENERAL INFORMATION:
; APPLICANT: Bushy, Sharon
; APPLICANT: Bushy, Sharon
; TITLE OF INVENTION: Transglutaminase Cross-Linkable
; POLYPEPTIDES AND METHODS RELATING THERETO
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East

CITY: Seattle
 STATE: WA
 COUNTRY: USA
 ZIP: 98102
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/483,236
 FILING DATE: 02-MAY-1998
 PRIORITY DATE: 03-MAY-1990
 ATTORNEY/AGENT INFORMATION:
 NAME: PARKER, Gary E.
 REGISTRATION NUMBER: 31-648
 REFERENCE/DOCKET NUMBER: 93-09c1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 206-442-6673
 TELEFAX: 206-442-6678
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5 amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Peptide
 US-08-483-236-15

Query Match 100.0%; Score 27; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred No. 1.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGV 5
 DB 1 VPGV 5

RESULT 13
 US-08-469-692-15
 Sequence 15, Application US/08469692
 Patent No. 5955035
 GENERAL INFORMATION:
 APPLICANT: Lees, Robert S. et al.
 TITLE OF INVENTION: SYNTHETIC PEPTIDES FOR ARTERIAL IMAGING
 NUMBER OF SEQUENCES: 27
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2804
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,692
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/498,046
 FILING DATE: 02-MAR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/201,057
 FILING DATE: 24-FEB-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/018,569
 FILING DATE: 01-MAR-1993
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/694,929
 FILING DATE: 02-MAY-1991
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,215
 FILING DATE: 03-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/518,142
 FILING DATE: 03-MAY-1990
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/189,130
 FILING DATE: 02-MAY-1988
 ATTORNEY/AGENT INFORMATION:
 NAME: Clark, Paul T.
 REGISTRATION NUMBER: 30,162
 REFERENCE/DOCKET NUMBER: 04547/002002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617/542-5070
 TELEFAX: 617/542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-469-692-15

Query Match 100.0%; Score 27; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred No. 1.4e+05;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGV 5
 DB 1 VPGV 5

RESULT 14
 US-08-913-364-6
 Sequence 6, Application US/08911364
 Patent No. 5959106
 GENERAL INFORMATION:
 APPLICANT: ROTHSTEIN, Aser
 APPLICANT: KEELY, Fred W.
 TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
 ELASTIN AND OTHER FIBROUS PROTEINS
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: FOLEY & LARDNER
 STREET: 1000 K Street, N.W.
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20007-5109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/911,364
 FILING DATE: 06-JUN-1997
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/023,552
 FILING DATE: 07-AUG-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Bent, Stephen A.
 REGISTRATION NUMBER: 29,768
 REFERENCE/DOCKET NUMBER: 041082/0104
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 672-5300

TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-911-364-6

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSVG 5
|||||
DB 1 VPSVG 5

RESULT 15
US-08-735-692-20
Sequence 20, Application US/08735692B
Patent No. 5972406
GENERAL INFORMATION:
APPLICANT: Utiy, Dan W.
APPLICANT: Sherry, Peter R.
APPLICANT: Prasad, Mari
INVENTOR: Utiy, Dan W.; Sherry, Peter R.; Prasad, Mari
FILE REFERENCE: BRF-011/01US
CURRENT APPLICATION NUMBER: US/08/735.692B
CURRENT FILING DATE: 1995-10-16
EARLIER APPLICATION NUMBER: 08/423,517
EARLIER FILING DATE: 1995-04-14
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 5
MOLECULE TYPE: peptide
FEATURE:
ORIGINISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-08-735-692-20

Query Match 100.0%; Score 27; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e-05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPSVG 5
|||||
DB 1 VPSVG 5

Search completed: April 24, 2001, 16:36:24
Job time: 304 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:42:01 ; Search time 74.56 seconds
(without alignments)
4.609 Million cell updates/sec

Title: US-09-340-736-6
Perfect score: 27
Sequence: 1 VPGVG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sequences: 198801 seqs, 58722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PR-67.*
1: P11.*
2: P12.*
3: P13.*
4: P14.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	27 100.0	38	2 T46593	phytoene dehydrog
2	27 100.0	76	2 I45885	elastin - bovine
3	27 100.0	93	2 C40585	hypothetical prote
4	27 100.0	120	2 A75586	hydrogenase expres
5	27 100.0	127	2 A72712	hypothetical prote
6	27 100.0	139	2 C70680	hypothetical prote
7	27 100.0	170	2 B83664	hypothetical prote
8	27 100.0	181	2 T05925	hypothetical prote
9	27 100.0	195	2 G64359	adenylate kinase
10	27 100.0	199	2 B95934	DNA polymerase, ba
11	27 100.0	204	2 B78292	probable amidotran
12	27 100.0	208	2 B78292	probable amidotran
13	27 100.0	208	2 T45249	probable two-comp
14	27 100.0	213	2 T36699	hypothetical prote
15	27 100.0	214	2 T23593	glutamine amidotra
16	27 100.0	222	2 J00640	hypothetical prote
17	27 100.0	231	2 D72532	hypothetical prote
18	27 100.0	246	2 T37169	hypothetical prote
19	27 100.0	255	2 A70778	hypothetical prote
20	27 100.0	271	2 S76871	hypothetical prote
21	27 100.0	274	2 G70898	probable 27F prot
22	27 100.0	274	2 B49930	orotidine-5'-phosp
23	27 100.0	274	2 B49930	hypothetical prote
24	27 100.0	283	2 B38694	hypothetical prote
25	27 100.0	285	2 S29306	poly(13-hydroxyalka
26	27 100.0	293	2 B75441	conserved hypothet
27	27 100.0	294	2 B59968	conserved hypothet
28	27 100.0	303	1 S75782	methanol dehydroge
29	27 100.0	305	2 B75308	amino acid ABC tra

ALIGNMENTS

RESULT 1
T46593
phytoene dehydrogenase [imported] - Mycobacterium marinum (fragment)
C:Species: Mycobacterium marinum
C:Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C:Accession: T46593
J. Ramakrishnan, L.; Tran, H.T.; Federspiel, N.A.; Falkow, S.
A:Title: A CRIB homolog essential for photochromogenicity in Mycobacterium marinum: i
A:Accession: T46593
A:Reference number: 243096; MUID:97440136
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-38 <RAM>
A:Cross-references: EMBL:U92075; NID:g1928930; PIDN:AA871427.1; PTD:g1928931
A:Experimental source: strain M
C:Genetics:
A:Gene: crtI

Query Match 100.0%; Score 27; DB 2; Length 38;
Best Local Similarity 100.0%; Field No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPGVG 5
Db 1 VPGVG 5
RESULT 2
I45885
elastin - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 13-Aug-1999
R. Accession: I45885
R. Reference number: 605-623, 1984
R. Title: Biology of disease: Elastin: Relation of protein and gene structure to disea
A:Reference number: I45885; MUID:85059254
A:Accession: I45885
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-76 <ROS>
A:Cross-references: GB:M31891; NID:g165008; PIDN:AA96416.1; PTD:g552319
C:Genetics: 20/1: 58/1
C:Superfamily: elastin
Query Match 100.0%; Score 27; DB 2; Length 76;
Best Local Similarity 100.0%; Field No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

hypothetical prote
hypothetical prote
methanol dehydroge
methanol dehydroge
fimbrial adhesin p
hypothetical prote
collagen srt-1 pre
hypothetical prote
hypothetical prote
probable regulator
38K protein - huma
hypothetical prote
steroidogenic acut
hypothetical prote
hypothetical prote
hypothetical prote

QY 1 VPGVG 5
|||||
DB 34 VPGVG 38

RESULT 3
Hypothetical protein (ntb 3' region) - Streptomyces cinnamonensis (fragment)
C:Species: Streptomyces cinnamonensis
C>Date: 03-May-1999 #sequence_revision 03-May-1999 #text_change 24-Feb-1995
C:Accession: G40595
R:Barth, A.; Leiser, A.; Robinson, J.A.
J. Bacteriol. 175, 3511-3519, 1993
A>Title: Cloning, sequencing, and expression of the gene encoding methylmalonyl-coenzyme
A:Reference number: A40595; MUID:93273720
A:Accession: G40595
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <R&>
A:Cross-references: GB:LI0054
C:Genetics:
A:Superfamily: 07G
Start codon: GTG

Query Match 100.0%; Score 27; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 41 VPGVG 45

RESULT 4
Hydrogenase expression/formation HypA-related protein - Deinococcus radiodurans (strain
A75586
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: A75586
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <R&>
A:Cross-references: GB:AE001863; GB:AE001825; NID:5640670; PIDN:AAF12463.1; PID:9646075
A:Experimental source: strain R1
C:Genetics:
A:Gene: DRA0316
A:Map position: 2
A:Superfamily: hydrogenase accessory protein

Query Match 100.0%; Score 27; DB 2; Length 120;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 67 VPGVG 71

RESULT 5
Hypothetical protein APE112 - Aeropyrum pernix (strain K1)
A7712
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: A7712

R:Kavaravasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funabashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero
A:Reference number: A72450; MUID:99310339
A:Accession: A7212
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <R&>
A:Cross-references: DB:AF000040; NID:95104188; PIDN:BA80097.1; PID:41043883; PID:9
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE112
C:Superfamily: Aeropyrum pernix hypothetical protein APE112

Query Match 100.0%; Score 27; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 19 VPGVG 23

RESULT 6
Hypothetical protein Rv2437 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
R:Cole, S.; Bosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Barris, D.; Gordon
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,
Rajendram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MUID:98295987
A:Accession: C70680
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-139 <C&>
A:Cross-references: GB:281451; GB:AL23456; NID:93261662; PIDN:CA803782.1; PID:625049
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv2437

Query Match 100.0%; Score 27; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 132 VPGVG 136

RESULT 7
Hypothetical protein BH014 [Imported] - Bacillus halodurans (strain C-125)
B33664
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, N.; Masui, F.; H
Nucleic Acids Res. 28, 4317-4331, 2000
A>Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
A:Reference number: A83650; MUID:20263314
A:Accession: B33664
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-170 <STO>
A:Cross-references: GB:AF001507; GB:BA000004; NID:910172612; PIDN:BA803833.1; GSPDB:G
A:Experimental source: strain C-125

C:Genetics:
A:Gene: BH0114

Query Match 100.0%; Score 27; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 58 VPGVG 62

RESULT 8
T05925
hypothetical protein - barley
C:Species: Hordeum vulgare (barley)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 08-Oct-1999
A:Accession: T05925
A:Residues: 1-181 <RES>
A:Molecule type: mRNA
A:Cross-references: EMBL:AJ222779; NID:el203989; PID:el203990
A:Experimental source: cv. Haisa, leaf

Query Match 100.0%; Score 27; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 120 VPGVG 124

RESULT 9
G64150
adenylate kinase (EC 2.7.4.3) M0479 [similarity] - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
A:Accession: G64359
A:Residues: 1-195 <BUL>
A:Molecule type: DNA
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Cross-references: GB:U67498; GP:L77117; NID:gl591180; PID:RAE98470.1; PID:gl591182;
A:Experimental source: strain K1

Query Match 100.0%; Score 27; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 14 VPGVG 18

RESULT 10
E69534
DNA polymerase, bacteriophage-type homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 08-Oct-1999
A:Accession: E69534
A:Residues: 1-199 <KLE>
A:Molecule type: DNA
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Cross-references: GB:AF000947; GB:AE000782; NID:g2468970; PID:RAE88977.1; PID:g2464
C:Superfamily: Archaeoglobus Probable DNA-Polymerase

Query Match 100.0%; Score 27; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 27 VPGVG 31

RESULT 11
E72695
adenylate kinase (EC 2.7.4.3) AP0981 [similarity] - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
A:Accession: E72695
A:Residues: 1-204 <KAW>
A:Molecule type: DNA
A:Status: preliminary
A:Cross-references: DDBJ:AF000060; NID:g5104188; PID:BAF79965.1; PID:g5104650
A:Experimental source: strain K1
C:Genetics:
A:Gene: AP0981
C:Superfamily: Sulfolobus adenylylate kinase
C:Keywords: phosphotransferase

Query Match 100.0%; Score 27; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 15 VPGVG 19

RESULT 12
D70544
probable amidotransferase - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
A:Accession: D70544
A:Residues: 1-195 <BUL>
A:Molecule type: DNA
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Cross-references: GB:U67498; GP:L77117; NID:gl591180; PID:RAE98470.1; PID:gl591182;
A:Experimental source: strain K1

Nature 393, 537-544, 1998
 A:Authors: Spates, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MIMD:98295987
 A:Accession: D70544
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-206 <COL>
 A:Cross-references: GB:Z95586; GB:AL13456; NID:G3261785; PID:CA809092.1; PID:G2117235
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: hsh
 C:Superfamily: amidotransferase hsh; trpG homology

Query Match 100.0%; Score 27; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 47 VPGVG 51

RESULT 13

T45249
 probable amidotransferase [imported] - *Mycobacterium leprae*
 C:Species: *Mycobacterium leprae*
 C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 04-Mar-2000
 A:Accession: T45249
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-206 <JAN>
 A:Cross-references: EMBL:AL049913; PID:CA843169.1
 A:Experimental source: cosmid B1610
 C:Genetics:
 A:Gene: hsh
 C:Superfamily: amidotransferase hsh; trpG homology

Query Match 100.0%; Score 27; DB 2; Length 206;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 47 VPGVG 51

RESULT 14

T36699
 probable two-component response regulator - *Streptomyces coelicolor* (fragment)
 C:Species: *Streptomyces coelicolor*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C:Accession: T36699
 A:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, April 1999
 A:Reference number: Z21597
 A:Accession: T36699
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 213 <MUR>
 A:Cross-references: EMBL:AL049731; PID:CA841738.1; GSPDB:GN00070; SC02DB:SCH66.11c
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCORDB:SCH66.11c
 C:Superfamily: cAMP protein; response regulator homology

Query Match 100.0%; Score 27; DB 2; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 192 VPGVG 196

RESULT 15

T23593
 hypothetical protein K10H10.4 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2000
 C:Accession: T23593
 R:Percy, C.

submitted to the EMBL Data Library, December 1996
 A:Reference number: Z19766
 A:Accession: T23593
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-214 <WIL>
 A:Cross-references: EMBL:Z83236; PID:CA805780.1; GSPDB:GN00020; CESE:K10H10.4
 A:Experimental source: clone K10H10
 C:Genetics:
 A:Gene: CESP:K10H10.4
 A:Map position: 2
 A:Mutrons: 11/3
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein K10H10.4

Query Match 100.0%; Score 27; DB 2; Length 214;
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 198 VPGVG 202

Search completed: April 24, 2001, 16:42:03
 Job time: 468 sec

Result	Query	Score	Length	DB	ID	Description
1	27	100.0	38	2	065422	005422 mycobacteri
2	27	100.0	76	2	028100	028100 bos taurus
3	27	100.0	120	2	028100	028100 bos taurus
4	27	100.0	126	2	028100	028100 bos taurus
5	27	100.0	127	1	095D00	095D00 aeropyrum p
6	27	100.0	130	6	018832	018832 equus caball
7	27	100.0	139	2	071592	071592 mycobacteri
8	27	100.0	141	2	065301	065301 unidentified
9	27	100.0	141	2	065301	065301 unidentified
10	27	100.0	141	2	065336	065336 unidentified
11	27	100.0	141	2	065364	065364 unidentified
12	27	100.0	141	2	065365	065365 unidentified
13	27	100.0	143	5	0592X5	0592X5 caenorhadi
14	27	100.0	162	10	09XIR0	09XIR0 arabidopsis
15	27	100.0	167	5	095V06	095V06 crotophila
16	27	100.0	179	3	095F00	095F00 aeropyrum lae
17	27	100.0	179	1	095F00	095F00 aeropyrum lae
18	27	100.0	176	11	05M3U3	05M3U3 mus muscula
19	27	100.0	181	4	095XK2	095XK2 homo sapien

ID	O97D00		PRELIMINARY;	PRT; 127 AA.
AC	Q97D00	1998	(T-EMBLrel_12, Created)	
DT	DI	01-NOV-1998	(T-EMBLrel_12, Last annotation update)	
DE	DG	01-MAY-2000	(T-EMBLrel_13, Last annotation update)	
DT	DI	01-MAY-2000	(T-EMBLrel_13, Last annotation update)	
DE	DG	SYNTHETICAL, 11.6 KDA PROTEIN APB1112.		
GN	APB1112			
OS	Aeropyrum pernix.			
CC	Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;			
OC	Cocci			
CC	Aeropyrum.			
OX	[1]	NCBJ_TaxID=56636;		
RN	SEQUENCE FROM N.A.			
RC	STRAIN:KI; 10339; PubMed=1032266;			
RL	MEDLINE=9511039; Hirata M., Sakiyama H., Yamazaki S., Halkawa Y.,			
RL	Takahashi T., Nakamura A., Takahashi K., Saito R., Iwano Baka S., Akai A., Kosugi H.,			
RA	Kakoyama A., Fukui S., Nagai Y., Mihiina K., Nakazawa H.,			
RA	Takanami M., Masuda S., Funabashi T., Tanaka T., Kudoh Y.,			
RA	Namurai J., Kusida N., Onuchi A., Aoki K., Kubota K., Nakamura Y.,			
RA	Nomura N., Sakoi Y., Kitachi H.;			
RT	"Complete genome sequence of an aerobic hyper-thermophilic			
RT	crenarchaeon. Aeropyrum pernix KI.";			
RL	J Biol Chem. 274(27):16290-16296, 1999.			
RL	DNA RES. 6:83-101(1999).			
RL	EMSL; AF000060; BAB0097.1;			
KW	HYPOTHETICAL protein.			
SQ	SEQUENCE 127 AA; 13575 MW; F4900D4ZBI09SD57 CRC64;			
<hr/>				
Query Match	100.0%; Score 27; DB 1; Length 127;			
Best Local Similarity	100.0%; Pred No. 2,7e+02;			
Matches	5; Conservative 0; Mismatches 0; Indels 0; Gaps			
QY	1 VFGVG 5			
DB	19 VFEGV 23			
RESULT	5			
ID	I18832		PRELIMINARY;	PRT; 130 AA.
DC	O18832			
DT	DI	01-NOV-1998	(T-EMBLrel_08, Created)	
DT	DI	01-NOV-1998	(T-EMBLrel_08, Last sequence update)	
DT	DI	01-MAY-2000	(T-EMBLrel_13, Last annotation update)	
DE	DG	AGGERCAN CORE PROTEIN (CARTILAGE-SPECIFIC PROTEOGLYCAN CORE PROTEIN)		
DE	DE	(CSGPC) (FRAGMENT).		
OS	Equis caballus (Horse).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Perissodactyla; Equidae; Equus.			
OX	[1]	NCBJ_TaxID=97996;		
RN	SEQUENCE FROM N.A.			
RC	STRAIN:CHONDROCYTES;			
RL	MEDLINE=98205637; PubMed=9550267;			
EX	Plowley C.P., Little C.B., Caenerson R.;			
RT	Molecular cloning and sequence analysis of the aggrecan interglobular			
RT	domain from porcine, equine, bovine and ovine cartilage: comparison of			
RT	proteinase-susceptible regions and sites of keratan sulfate			

[illegible]

CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS

CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -!- PTH: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
CC -!- CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE AGGREGAN/VENUSCAN PROTEOGLYCAN FAMILY.
CC ENBL_TaxID=1773; NCBI_TaxID=1773;
DR INTERPRO: IPR000539; -!
DR PROSITE: PS01241; LINK: PARTIAL.
KW Glycoprotein; Cartilage; Proteoglycan; Repeat.
FT NON_TER 1 1
FT DOMAIN <1 10 LINK 2.
FT DOMAIN <1 10 G1-B'.
FT CARBOHYD 48 48 POTENTIAL.
FT CARBOHYD 95 95 POTENTIAL.
FT NON_TER 130 130
FT NON_TER 130 130
SQ SEQUENCE 130 AA; 13930 MW; 082F79B5AE79B53E CRC64;
Query Match 100.0%; Score 27; DB 6; Length 130;
Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
DB 82 VPGVG 86
RESULT 6
PRT: 139 AA.
AC P71912 PRELIMINARY;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE HYPOHETICAL 15.3 KDA PROTEIN.
GN RV2437 OR MFCY428.09C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxID=1773;
FT NON_TER 1 1
RP SEQUENCE FROM N.A.
RC MEDLINE=H37RV;
EX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
Gordon S.V., Eigleier K., Gas S., Barry C.E. III, Tekala F.,
Badoock K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Hornsbly T., Jegels K., Krogh A., McLean J., Moule S., Murphy L.,
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
Taylor K., Whitehead S., Barrall B.C.;
RA T. (2000) Complete genome sequence of Mycobacterium tuberculosis from the
Haitian isolate H37Rv.
RL Nature 393:537-544(1998).
DR ENBL; 261451; CAB03782.1; -!
DR TUBERCULIST; RV2437; -!
KW Hypothetical protein.
SQ SEQUENCE 139 AA; 15315 MW; 842BF115C0E102EC CRC64;

Query Match 100.0%; Score 27; DB 2; Length 139;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
DB 132 VPGVG 136
RESULT 7
O66301

ID O66301 PRELIMINARY; PRT: 141 AA.
AC O66301;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
CS UNIDENTIFIED NITROGEN-FIXING BACTERIA.
OC Bacteria.
CX NCBI_TaxID=34107;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohkuma M., Noda S., Kudo T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EXBL; AB011855; BAA28392.1; -!
DR HSP; P00456; 1CF2.
DR INTERPRO: IPR000392; -!
DR PFAM; PF00142; fer4_Nifs; 1.
DR PRINTS; PR00091; NITROGNASEII.
DR PROSITE; PS00692; NIFH_FRXC_2; 1.
DR PROSITE; PS00746; NIFH_FRXC_1; 1.
FT NON_TER 1 1
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 14882 MW; 024548CD55EDBF40 CRC64;
Query Match 100.0%; Score 27; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
DB 74 VPGVG 78
RESULT 8
O66330 PRELIMINARY; PRT: 141 AA.
AC O66330;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
CS UNIDENTIFIED NITROGEN-FIXING BACTERIA.
OC Bacteria.
CX NCBI_TaxID=34107;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohkuma M., Noda S., Kudo T.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EXBL; AB011900; BAA28435.1; -!
DR HSP; P00456; 1CF2.
DR INTERPRO: IPR000392; -!
DR PFAM; PF00142; fer4_Nifs; 1.
DR PRINTS; PR00091; NITROGNASEII.
DR PROSITE; PS00692; NIFH_FRXC_2; 1.
DR PROSITE; PS00746; NIFH_FRXC_1; 1.
FT NON_TER 1 1
FT NON_TER 141 141
SQ SEQUENCE 141 AA; 14632 MW; 7251716CF65C6732 CRC64;
Query Match 100.0%; Score 27; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 3.1e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
DB 74 VPGVG 78
RESULT 9

O66336
 ID O66336 PRELIMINARY; PRT: 141 AA.
 AC O66336
 DT 01-AUG-1998 (TRENDEL. 07, Created)
 DT 01-AUG-1998 (TRENDEL. 07, Last sequence update)
 DT 01-AUG-2000 (TRENDEL. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 CX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB011956; BAA28481.1;
 DR HSSP: P00456; 1CP2.
 DR PFAM: PF00142; fer4_NifH; 1.
 DR PRISM: PR00091; NITROGENASEII.
 DR PROSITE: PS00692; NIFH_FRXC_2; 1.
 DR PROSITE: PS00746; NIFH_FRXC_1; 1.
 FT NON_TER 1 1
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 14910 MW; D2E05148699D28C3 CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3; le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |||||
 DB 74 VPGVG 78

RESULT 10
 O66336 PRELIMINARY; PRT: 141 AA.
 AC O66336
 DT 01-AUG-1998 (TRENDEL. 07, Created)
 DT 01-AUG-1998 (TRENDEL. 07, Last sequence update)
 DT 01-MAY-2000 (TRENDEL. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 CX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB011954; BAA28487.1;
 DR HSSP: P00456; 1CP2.
 DR INTERPRO: IPR000392;
 DR PFAM: PF00142; fer4_NifH; 1.
 DR PROSITE: PS00692; NIFH_FRXC_2; 1.
 DR PROSITE: PS00746; NIFH_FRXC_1; 1.
 FT NON_TER 1 1
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 14780 MW; E069FA235F62484B CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3; le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |||||
 DB 74 VPGVG 78

RESULT 11

O66365
 ID O66365 PRELIMINARY; PRT: 141 AA.
 AC O66365
 DT 01-AUG-1998 (TRENDEL. 07, Created)
 DT 01-AUG-1998 (TRENDEL. 07, Last sequence update)
 DT 01-MAY-2000 (TRENDEL. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 CX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB011955; BAA28488.1;
 DR HSSP: P00456; 1CP2.
 DR PFAM: PF00142; fer4_NifH; 1.
 DR PRISM: PR00091; NITROGENASEII.
 DR PROSITE: PS00692; NIFH_FRXC_2; 1.
 DR PROSITE: PS00746; NIFH_FRXC_1; 1.
 FT NON_TER 1 1
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 14786 MW; 2EC10A2BF3C05E2E CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3; le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |||||
 DB 74 VPGVG 78

RESULT 12
 O66367 PRELIMINARY; PRT: 141 AA.
 AC O66367
 DT 01-AUG-1998 (TRENDEL. 07, Created)
 DT 01-AUG-1998 (TRENDEL. 07, Last sequence update)
 DT 01-MAY-2000 (TRENDEL. 13, Last annotation update)
 DE ALPHA SUBUNIT OF DINITROGENASE REDUCTASE (FE PROTEIN) (FRAGMENT).
 GN NIFH.
 OS unidentified nitrogen-fixing bacteria.
 OC Bacteria.
 CX NCBI_TaxID=34107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Okuma M., Noda S., Kudo T.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB011957; BAA28490.1;
 DR HSSP: P00456; 1CP2.
 DR INTERPRO: IPR000392;
 DR PFAM: PF00142; fer4_NifH; 1.
 DR PROSITE: PS00692; NIFH_FRXC_2; 1.
 DR PROSITE: PS00746; NIFH_FRXC_1; 1.
 FT NON_TER 1 1
 FT NON_TER 141 141
 SQ SEQUENCE 141 AA; 14505 MW; 7A2B008BE56AC29C CRC64;

Query Match 100.0%; Score 27; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 3; le+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5
 |||||
 DB 74 VPGVG 78

RESULT 13

Q9N2X5

```

Query Match      100.0%; Score 27; DB 5; Length 167;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VPGVG 5
Db       52 VPGVG 56
      |||||

```

Search completed: April 24, 2001, 16:40:37
Job time: 427 sec

CC Linking component. The copolymer is useful in prosthetic systems, for
 CC repairing a natural elastic system. It is so functionalised so as to
 CC provide reactive gps. which can become covalently cross-linked by
 CC tissue enzymes to newly synthesised connective tissue protein.
 CC The copolymer comprises 15 units of VPQG, 5 units of the block
 CC unit and 1 unit of ABAABABACA. It has a mol. wt. of 80000.
 XX
 SQ Sequence: 4 AA;

Query Match 100.0%; Score 23; DB 7; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQG 4
 Db 1 VPQG 4

RESULT 2

R65231 R65231 standard; peptide; 4 AA.

XX R65231;

DT 12-OCT-1995 (first entry)

DE Fibrinogen/beta-casein transglutaminase cross-linkable copolymer.

XX Transglutaminase cross-linkable copolymer; fibrinogen; beta-casein;

KW biocompatible; bioadhesive polymers; wound repair; prosthetics;

XX bone and soft tissue matrices; controlled drug release carriers.

XX Synthetic.

XX WO9505396-A.

XX 23-FEB-1995.

XX 05-AUG-1994; 94WO-US08754.

XX 13-AUG-1993; 93US-0106509.

XX (ZYMO) ZYMOGENETICS INC.

XX Busby SJ, Labroc VM;

XX WPI; 1995-098722/13.

XX New peptide(s) based on fibrinogen and beta-casein -

XX cross-linkable by transglutaminase, used for preparing

XX biocompatible, bioadhesive polymers

XX Claim 16; Page 48; 58pp; English.

XX A biocompatible, bioadhesive, transglutaminase cross-linkable

XX copolymer comprises a first polypeptide monomer from 13-120 amino

XX acids containing a segment of the formula S1-Y-S2 cross-linkable

XX by a transglutaminase, where S1 = R65218, S2 = R65219 and Y = a

XX spacer of 0-7 amino acids pref. R65220/21/22 or R65223; and

XX a second polypeptide monomer selected from R65228-R65231 (the

XX first monomer can also be R65227). The copolymer can be used to

XX produce tissue adhesives, wound repair formulations, rigid

XX prosthetics, matrices for the replacement of bone and soft tissue

XX structures and carriers for controlled drug release compans.

XX Sequence 4 AA;

Query Match 100.0%; Score 23; DB 16; Length 4;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQG 4
 Db 1 VPQG 4

RESULT 3

R18264 R18264 standard; peptide; 4 AA.

XX R18264;

XX 29-AUG-1997 (first entry)

DE Transglutaminase cross-linkable polypeptide elastomeric peptide.

XX Elastomeric; homopolymer; copolymer; tissue sealant; skin graft;

XX burn.

XX Synthetic.

XX NO9640780-A1.

XX 19-DEC-1996.

XX 31-MAY-1996; 96NO-US08269.

XX 07-JUN-1995; 95US-0483236.

XX (ZYMO) ZYMOGENETICS INC.

XX Busby SJ, Labroc VM;

XX WPI; 1997-052237/05.

XX Transglutaminase cross-linkable peptide(s) - used in the mfr. of

XX biocompatible, bioadhesive tissue sealant and wound healing

XX preparations.

XX Claim 6; Page 48; 55pp; English.

XX Novel polypeptides optionally having one or both of the amino-terminus

XX and carboxy-terminus flanked by an elastomeric peptide, which is cross-

XX linkable by a transglutaminase comprises a segment of formula:

XX S1-Y-S2. The present sequence represents a specifically claimed

XX example of an elastomeric peptide. The homo- and copolymers produced

XX are useful in tissue sealant and wound healing formulations. Tissue

XX sealants are useful in skin grafting for burn victims and for sealing

XX surgical and other wounds.

XX Sequence 4 AA;

Query Match 100.0%; Score 23; DB 18; Length 4;

Best Local Similarity 100.0%; Pred. No. 3.2e+05;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQG 4

Db 1 VPQG 4

RESULT 4

Y31685 Y31685 standard; peptide; 4 AA.

XX Y31685;

XX 22-NOV-1999 (first entry)

DE Tetrapeptide used in novel elastomer polymers.

KW Elastomer; bioelastomer; polymer; tissue augmentation;

XX tissue restoration; tissue reconstruction; tissue repair; implant.

XX Synthetic.
 US WO9943271-A1.
 PA
 XX
 XX
 PD 02-SEP-1999.
 XX
 XX 26-FEB-1999; 99WO-US04440.
 XX
 XX 29-MAY-1998; 98US-0087155.
 XX 27-FEB-1998; 98US-0076297.
 XX (BIOE-) BIOELASTICS RES LTD.
 PA
 XX Glazer PA, Parker TM, Urry DM;
 PI
 XX WPI; 1999-510487/45.
 DR
 CC Augmentation or restoration of mammalian tissue by injecting
 CC solution of peptide polymer used for soft or hard tissue
 CC reconstruction, especially of intervertebral disks
 PT
 XX claim 8; Page 75; 13pp; English.
 XX
 CC The present sequence represents an example of a tetrapeptide that
 CC is used in novel bioelastic polymers. The invention provides a
 CC method of tissue augmentation by injecting a polymer comprising
 CC repeated tetrapeptide units. The polymer is composed of monomeric
 CC tetrapeptide and tetrapeptide monomeric units where the monomeric
 CC units form a series of beta-turns separated by dynamic bridging
 CC segments. The polymer has an inverse temperature transition value
 CC that is less than the tissue temp. and is injected in water
 CC solution at concavate concentration. The polymer can be injected
 CC at perineural or subdural sites (for treatment of urinary
 CC incontinence or for cosmetic purposes), or into hard or soft
 CC tissue. The polymer can be used for tissue specific
 CC application in restoration of intervertebral discs.
 XX
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 23; DB 20; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.2e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 VPGG 4
 1 VPGG 4
 RESULT 5
 B19213
 XX B19213 standard; Peptide: 4 AA.
 AC B19213;
 XX
 DT 19-FEB-2001 (first entry)
 DE Elastin fragment used to make hybrid bay scallop abductions.
 XX
 XX Bay scallop; abductin; chemomechanical transduction; drug delivery;
 XX inverse temperature transition; water soluble drug; biomaterial;
 XX fabric; organ prostheses.
 KW
 XX Synthetic.
 OS
 XX US6127166-A.
 PN
 XX 03-OCT-2000.
 PD
 XX 03-NOV-1997; 97US-0963168.
 XX
 XX 03-NOV-1997; 97US-0963168.

XX (BAYL/) BAYLEY H.
 PA (BAYL/) BAYLEY H.
 PA (WANG/) WANG Y.
 XX
 XX Bayley H, Cao Q, Wang Y;
 PI
 XX WPI; 2000-611057/58.
 DR
 XX Abductin nucleic acid molecules, useful for expressing abductin
 XX polypeptides and methods of administering the same.
 PT
 XX Vehicles for administering water soluble drugs.
 XX
 PS Disclosure; Column 14; 30pp; English.
 XX
 CC The present sequence is used to make hybrid bay scallop abductin
 CC polypeptide sequences. Abductin is capable of chemomechanical
 CC transduction or inverse temperature transition. Abductin polypeptides
 CC are used to make hybrid bay scallop abductin polypeptides in
 CC the manufacture of drug delivery vehicles for administering water
 CC soluble drugs. The abductin polypeptides and their derivatives are
 CC also useful in the manufacture of broad range of biomaterials ranging
 CC from light-weight durable fabric for clothing to matrices useful for
 CC human tissue and organ prostheses.
 XX
 SQ Sequence 4 AA;
 Query Match 100.0%; Score 23; DB 21; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGG 4
 DB 1 VPGG 4
 RESULT 6
 Y80336
 ID Y80336 standard; peptide; 4 AA.
 XX
 XX Y80336;
 AC
 XX 24-MAY-2000 (first entry)
 XX
 DE Repeat elastin tetrapeptide from tropoelastin.
 KW
 KW Overexpression; bioelastic polypeptide; pentapeptide; tetrapeptide;
 KW hexapeptide; nonapeptide; repeat unit; inclusion body; elastin;
 KW inverse temperature transition; tropoelastin.
 XX
 XX Bos taurus.
 XX
 XX US6004782-A.
 PN
 XX 21-DEC-1999.
 PD
 XX 13-OCT-1995; 95US-0542051.
 PF
 XX 14-APR-1995; 95US-0423642.
 PR
 XX (BIOE-) BIOELASTICS RES LTD.
 XX (BAYR-) BAY RES FOUND.
 PA
 XX Xu J, Daniell H, McPherson DT, Urry DM;
 PI
 XX WPI; 2000-125738/11.
 DR
 XX Overexpression of bioelastic polypeptides, in prokaryotic cells,
 XX exhibiting an inverse temperature transition
 PT
 XX Disclosure; Column 2; 32pp; English.
 XX

CC The invention relates to overexpressing a bioelastic polypeptide in
 CC a prokaryotic host cell by introducing into the host cell a vector
 CC containing a promoter operably linked to a nucleic acid encoding a
 CC bioelastic polypeptide, and growing the host cell for expressing
 CC the polypeptide. The bioelastic polypeptide especially comprises
 CC pentapeptide, tetrapeptide, hexapeptide, or nonapeptide repeat units.
 CC The host cell is grown such that the overexpressed peptide is
 CC produced in inclusion bodies, where the volume of inclusion bodies
 CC comprises 0-50% (v/v) of the total cellular volume of the host cell
 CC and the bioelastic polypeptides of polymers
 CC that exhibit an inverse temperature transition are
 CC represents the natural elastin tetrapeptide repeat from tropoelastin.

SQ Sequence 4 AA;

Query Match 100.0%; Score 23; DB 21; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 Db 1 VPSS 4

RESULT 7
 ID B63971 standard; Peptide; 4 AA.
 AC B63971;
 DT 19-MAR-2001 (first entry)
 DE Elastin repeating unit peptide sequence SEQ ID 1.
 KW Proteinaceous polymer; repeat unit; structural polymer; coating; film;
 KW fibre; membrane; adhesive; emulsion; laminate; keratin; collagen.
 OS Unidentified.
 PN US6140072-A.
 PD 31-OCT-2000.
 PF 07-JUN-1995; 95US-0475411.
 PR 06-NOV-1990; 90US-0609716.
 PR 24-OCT-1997; 97US-0727258.
 PR 24-OCT-1997; 97US-0727258.
 PR 09-NOV-1998; 88US-0259429.
 XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
 PI Cappello J, Ferrari FA;
 DR WPI; 2001-048958/06.
 XX New DNA encoding a polymer with strands of repeating units of natural
 PT protein joined by intervening oligopeptide for producing high molecular
 PT weight polymers of amino acids
 XX Claim 3; Column 143; 73pp; English.
 CC This invention relates to DNA encoding a proteinaceous polymer. The
 CC polymer comprises strands of repeating units of a natural protein capable
 CC of assembling into aligned structures, with at least 2 strands joined by
 CC an intervening oligopeptide other than the repeating units. The
 CC intervening oligopeptide is unaligned and the polymer has individual
 CC strands of the same or different repeating units. The DNA is useful for
 CC producing high molecular weight polymers of amino acids based on
 CC biologically and chemically active structural polymers. These polymers
 CC may be used to provide a variety of structures for different purposes,
 CC and to produce articles including coatings, or other (non)structural

CC components, e.g. fibres, films, membranes, adhesives or emulsions, or
 CC with other compounds and/or compositions to form composites or laminates.
 CC Peptide sequences B63971-B63991 represent monomer sequences which can be
 CC used in the polymers of the invention. Oligonucleotide sequences
 CC F23370 - F23386 and amino acid sequences B63992 - B64002 are used in the
 CC construction of SLP and PCB-SLP polymers. Oligonucleotide sequences
 CC F23387 - F23397 and amino acid sequences B64003 - B64008 are used in the
 CC construction of CLP (collagen like protein) polymers. Oligonucleotide
 CC sequences F23398 - F23409 and amino acid sequences B64009 - B64014 are
 CC used in the construction of Keratin polymers. Proteins and peptides
 CC represented by sequences B64015 - B64049 are examples of polymers of the
 CC invention.

SQ Sequence 4 AA;

Query Match 100.0%; Score 23; DB 22; Length 4;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 Db 1 VPSS 4

RESULT 8
 ID W07157 standard; peptide; 7 AA.
 AC W07157;
 DT 24-JAN-1997 (first entry)
 DE Synthetic peptide used in GalNac-transferase activity SPA.
 KW SPA, scintillation proximity assay; antigen; bead coating; capture;
 KW antibody; N-acetyl galactosamine transferase; GalNac transferase;
 KW activity; enzyme; O-linked glycosylation.
 OS Synthetic.
 XX W09615258-A1.
 XX 23-MAY-1996.
 XX 08-NOV-1995; 95WO-US13483.
 XX 16-NOV-1994; 94US-0340283.
 XX (UPJO) UPJOHN CO.
 XX Elhammer AP;
 XX WFI; 1996-268220/27.
 XX Scintillation proximity assay for N-acetyl-galactosaminyl activity
 PT - esp. for large scale screening of cpds. for their effect on enzyme
 PT activity
 XX Claim 14; Page 17; 29pp; English.
 XX W06985-W07180 are antigenic peptides derived from either the
 CC the human or the mouse flag peptide (FLAG peptide (Kodak)).
 CC The peptides are used in the assay for N-acetyl-galactosaminyl transferase (GNT).
 CC proximity assay for N-acetyl-galactosamine (GalNAc) transferase (GNT)
 CC activity. The assay involves fewer steps than known assays and is
 CC quicker, producing excellent signal-to-noise ratios. The
 CC assay is capable of screening large numbers of cpds. for their
 CC ability to affect GNT activity and is thus useful for identifying
 CC inhibitors and promoters of glycosylation (in partic. O-linked
 CC glycosylation).

SQ Sequence 7 AA;

Query Match 100.0%; Score 23; DB 17; Length 7;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGG 4
 Db 4 VPGG 7

RESULT 9

W59326
 ID W59326 standard; peptide; 7 AA.

XX WC
 XX WC
 XX WC

24-SEP-1998 (first entry)

Non-polio enterovirus peptide fragment 61S.

XX Non-polio enterovirus; NPEV; enteroviral disease; aseptic meningitis;
 KW vaccination.

XX Enterovirus sp.

XX W09814611-A2.

XX 09-APR-1998.

01-OCT-1997; 97MO-US17734.

02-OCT-1996; 96US-0027353.

(USSH) US DEPT HEALTH & HUMAN SERVICES.

Kilpatrick D;

WPI; 1998-240106/21.

Identifying non-polio enteroviruses - using primers which hybridise
 to same and antisense strands that encode conserved non-polio
 enterovirus peptide sequences

Claim 4; Page 23; 47pp; English.

The peptide sequences W59298-W59344 are amplified by primers to detect
 the presence of a non-polio enterovirus (NPEV) in a sample. The primers
 and assays are used to detect NPEVs in a sample, to serotype these
 viruses, to diagnose enteroviral diseases and medical conditions, and to
 correlate (or disprove a correlation between) specific symptoms or
 combinations of symptoms with the presence of a particular enterovirus.
 They can be used for diseases such as aseptic meningitis. The detection
 of NPEV infections and their correlation with medical conditions will
 make possible vaccines and methods of treatment.

Sequence 7 AA;

Query Match 100.0%; Score 23; DB 19; Length 7;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGG 4
 Db 3 VPGG 5

RESULT 10

Y50084
 ID Y50084 standard; peptide; 7 AA.

XX Y50084;

XX 19-JAN-2000 (first entry)
 DT Coxsackievirus A VPI conserved epitope 61.
 DE Virus; epitope; target; degenerate; PCR; primer; amplification;
 KW V21; nonstructural protein 2A; conserved; base analogue; inosine;
 KW Predetermined nucleotide; diagnosis; enterovirus; poliovirus.
 XX Synthetic.
 OS Coxsackievirus.
 XX W09953097-A2.
 PN 21-OCT-1999.
 PD 06-APR-1999; 99MO-US07513.
 XX 15-APR-1998; 98US-0081944.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Kilpatrick DR;
 PI WPI; 1999-620444/53.
 DR N-F808; 232603.
 DR Designing degenerate polymerase chain reaction primers -
 Example 3; Page 18; 30pp; English.
 XX This sequence represents a conserved Coxsackievirus A
 CC (serotype 21) VPI epitope. The invention relates to a novel
 CC method for designing degenerate PCR primers (230975-231000, 232601-
 CC 232611) for amplifying target polynucleotides. This method comprises
 CC identifying uniquely conserved amino acid sequences (e.g., this
 CC epitope) in target proteins; synthesizing degenerate polynucleotides
 CC encoding the conserved sequences; and substituting the synthesized
 CC polynucleotides with up to four predetermined nucleotides (e.g.,
 CC inosine) at degenerate nucleotide positions. The nucleic acids
 CC comprise no more than 7 degenerate positions, have no more
 CC than adjacent predetermined nucleotides and the predetermined
 CC strand. The degenerate primers are useful for amplifying target polynucleotides
 CC by the polymerase chain reaction (PCR). The use of the method of
 CC designing degenerate primers useful for the detection of polioviruses
 CC in clinical samples is described in US585477. The degenerate primers
 CC facilitate PCR amplification of unknown polynucleotides, where the amino
 CC acid sequence encoded is known. The primers also allow for the
 CC correlation of the subsequent molecular based diagnosis with a
 CC serologically derived diagnosis.
 XX Sequence 7 AA;

Query Match 100.0%; Score 23; DB 20; Length 7;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPGG 4
 Db 3 VPGG 6

RESULT 11

W07039
 ID W07039 standard; peptide; 9 AA.

XX W07039;

17-JAN-1997 (first entry)

Synthetic peptide used in GalNac-transferase activity SPA.

XX SP4: scintillation proximity assay; antigen; bead coating; capture;
 KW antibody; N-acetyl galactosamine transferase; Galnac transferase;
 KW activity; enzyme; O-linked glycosylation.
 XX Synthetic.
 XX WO9515258-A1.
 XX 23-MAY-1996.
 XX 08-NOV-1995; 95WO-US3483.
 XX 16-NOV-1994; 94US-0340283.
 XX (UPTO) UPTOHN CO.
 XX Elhammer AP;
 XX WPI; 1996-268220/27.
 XX Scintillation proximity assay for N-acetyl:galactosaminyl activity
 - esp. for large scale screening of cpds. for their effect on enzyme
 activity
 XX Claim 14; Page 17; 29pp; English.
 XX WO6985-007180 are antigenic peptides derived from either the
 CC product of the human c-myc oncogene or the FLAG peptide (DYKDDK).
 CC The peptides are useful for coating beads used in a scintillation
 CC proximity assay for N-acetyl:galactosamine (Galnac)-transferase (GNT)
 CC activity. The assay involves fewer steps than known assays and is
 CC quicker, producing excellent signal-to-noise ratios. The
 CC assay is capable of screening large numbers of cpds. for their
 CC ability to affect GNT activity and is thus useful for identifying
 CC inhibitors and promoters of glycosylation (in partic. O-linked
 CC glycosylation).
 XX Sequence 9 AA;

Query Match 100.0%; Score 23; DB 17; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3.2e+05;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VP6G 4
 DB 1 VP6G 4

RESULT 12

ID P91309 standard; Protein; 10 AA.
 XX AC P91309;
 DT 09-MAR-1992 (first entry)
 XX Sequence of beta-turn of a bioelastomeric material.
 XX Bioelastomer: elastomeric material; wound dressing; burn;
 KW artificial veins; arteries; skin; ligament; biodegradable.
 XX Key Location/Qualifiers
 XX Misc-difference 1..3
 FT /note= "May be PG, GG, G or a covalent
 bond"
 FT Misc-difference 8..10
 FT /note= "May be VPG, VP, V or a covalent
 bond"
 FT Misc-difference 4..7
 FT /note= "May be absent or present up to 200
 times"

XX WO8910099-A.
 PN 02-NOV-1989.
 XX 14-APR-1989; 89WO-US01482.
 XX 21-APR-1988; 88US-0184407.
 XX (UABR-) UAB-RES FOUNDATION.
 XX Dirty DW;
 XX WPI; 1989-339745/46.
 XX Elastomeric polypeptide material - a useful for preventing
 PN adhesion between tissues and wound repair sites
 XX Claim 8; page 86; 93pp; English.
 XX The elastomeric material of the invention comprises a bioelastomer
 CC contg. repeating elastomeric tetrapeptide or pentapeptide units opt.
 CC modified by hexapeptide units, the units consisting of hydrophobic
 CC AA residues and Gly residues and existing in a conformation having
 CC a beta-turn. The beta-turn of the the bioelastomer comprises
 CC a polypeptide unit of VPGVG (see P91307) and/or IPGVGVPXVG
 CC (see P91308), or a pentapeptide repeating unit of formula
 CC (see P91309), or a polypeptide unit of formula P91310, and the hexapeptide
 CC IPGVG (see P91310).
 XX repeating unit of formula ARGVG (see P91313).
 XX Sequence 10 AA;

Query Match 100.0%; Score 23; DB 10; Length 10;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VP6G 4
 DB 4 VP6G 7

RESULT 13

ID W02011 standard; peptide; 10 AA.
 XX AC W02011;
 DT 30-SEP-1996 (first entry)
 XX Peptide derived from hepatitis C virus NS4 protein.
 XX Antigenic peptide; hepatitis C virus; HCV; non-A non-B;
 KW non-structural protein 4; non-reactive; NS4;
 XX Hepatitis C virus.
 XX WO9604300-A1.
 PD 15-FEB-1996.
 XX 28-JUL-1995; 95WO-US09599.
 XX 29-JUL-1994; 94US-0282758.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX (USSH) US SEC DEPT HEALTH.
 XX Fields HA, Khudayakov YE;
 XX WPI; 1996-129330/13.

PF Antigenic peptide(s) binding anti-hepatitis C virus antibodies -
 PT useful for differential diagnosis of HCV in subjects

XX Example: Page 39; 50bp; English.

CC The present peptide, which was derived from the hepatitis C virus
 CC (HCV) non-structural protein 4 (NS4), was prep'd. using Fmoc
 CC chemistry. It was tested for reactivity against a panel of 32
 CC anti-HCV positive sera, and was found to react with 15 of them.

SQ Sequence 10 AA;

Query Match 100.0%; Score 23; DB 17; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VPQG 4
 |||||
 7 VPQG 10

RESULT 14

W41471

ID W41471 standard; peptide: 10 AA.

XX AC W41471;

XX AC W41471;

XX 05-JUN-1998 (first entry)

Fragment of nematode alpha 3-tubulin.

XX Nematode: alpha 3-tubulin; aligned peptide array: protein binding site;

XX protein interaction site; ligand detection.

XX Nematoda.

XX EP818467-A2.

XX 14-JAN-1998.

XX 11-JUL-1997; 97EP-0111868.

XX 12-JUL-1996; 96JP-0183140.

XX (NIDE) NEC CORP.

XX Miwa J, Siddiqui SS;

XX WPI; 1998-065262/07.

XX Aligned peptide array - for detecting protein binding or interaction
 XX sites, etc.
 XX Example 1; Page 5; 27bp; English.

CC This sequence represents a fragment of alpha 3-tubulin derived from a
 CC nematode. This sequence was used as a part of an aligned peptide array of
 CC the invention. The aligned peptide array comprises, as separate elements,
 CC peptide segments obtained by dividing the amino acid sequence of a
 CC protein into sequence segments of any suitable length and of any suitable
 CC frame and synthesizing peptides on the basis of the sequence segments.
 CC preferably with the amino acid sequences of the peptide segments
 CC expressing the amino acid sequences of the protein. The peptide array can
 CC be used in a method for the detection of a binding or interaction site of
 CC a protein or the detection of a binding or interaction site of a protein in a
 CC method for the modification of a protein and a method for the design
 CC of a protein. Using the information from the above methods, ligands
 CC detected by them can be modified, or designed. The peptide array can also
 CC be used in enzyme immunoassay method, represented by ELISA. The array can
 CC be used in a viroimmunoassay method, a metalloimmunoassay method, a
 CC fluoroimmunoassay method or, a radioimmunoassay method.

SQ Sequence 10 AA;

Query Match 100.0%; Score 23; DB 19; Length 10;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VPQG 4

|||||

Db 5 VPQG 8

RESULT 15

Y31331

ID Y31331 standard; peptide: 11 AA.

XX AC Y31331;

XX 01-NOV-1999 (first entry)

B. subtilis surface binding peptide ligand.

XX Bacterial spore; bioassaying: phase-display library; Bacillus; vaccine;

XX pathogen detection; biological warfare agent; B. anthracis.

XX Bacillus subtilis.

XX WO9536081-A1.

XX 22-JUL-1999.

XX 14-JAN-1999; 99WO-0500771.

XX 14-JAN-1998; 98US-0071411.

XX (UABR-) UAB RES FOUND.

XX Turnbough CL;

XX WPI; 1999-468943/39.

XX Identifying peptides that bind to the surface of bacterial spores by
 XX bioassaying phase-display library, useful as vaccines and diagnostic
 XX agents
 XX Disclosure: Page 9; 23pp; English.

XX The invention provides peptides that bind to the surface of bacterial
 XX spores. These peptides are identified by bioassaying a phase-display
 XX library with the spores. The library is incubated with spores and any
 XX phase-spore complexes formed recovered by centrifuging. They are washed
 XX thoroughly, then phase eluted with buffer, the eluate neutralised and
 XX eluted phase amplified. The procedure is repeated, for 3 or 4 rounds of
 XX panning, then individual clones purified, amplified and genomic DNA
 XX extracted for determination of peptide-encoding sequences. Peptides are
 XX tested for their ability to bind to the surface of bacterial spores.
 XX The peptide sequence used for capture and identification of bacterial
 XX spores, particularly of the genus Bacillus, particularly for detecting
 XX pathogens, or otherwise harmful species, in the environment (e.g. air,
 XX water or food) or in clinical samples. They are also used to protect
 XX against disease-causing spores (e.g. by incorporation in protective
 XX masks), and in vaccines to generate a protective antibody response. The
 XX peptides-specific antibodies are useful as detection reagents, e.g. in
 XX enzyme-linked immunosorbent assay. The peptides are species specific,
 XX i.e. they can differentiate between the possible biological warfare agent
 XX likely to be present in the environment and the present sequence
 XX represents a peptide that can bind to B. subtilis.

SQ Sequence 11 AA;

Query Match

100.0%; Score 23; DB 20; Length 11;

Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 UPSC 4
DB 5 WPSG 9

Search completed: April 24, 2001, 16:38:20
Job time: 420 sec

GenCore version 4.5
Copyright (C) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: April 24, 2001, 16:42:49 ; Search time 44.88 Seconds
(without alignments)
3.816 Million cell updates/sec

Title: US-09-340-736-6
Perfect score: 27
Sequence: 1 VPGVG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 93435 seqs, 3425486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	27	100.0	93	1 YMG3_SPCRM	Q05072 streptomyces
2	27	100.0	128	1 YMG3_CHEHL	Q26709 caenorhabdi
3	27	100.0	128	1 YMG3_CHEHL	Q26709 caenorhabdi
4	27	100.0	132	1 KADA_MRTGA	P43408 methanococ
5	27	100.0	132	1 KADA_MRTGA	P43409 methanococ
6	27	100.0	132	1 KADA_MRTGA	P43410 methanococ
7	27	100.0	132	1 KADA_MRTGA	P43411 methanococ
8	27	100.0	204	1 KADA_MRTGA	Q97442 aeropyrum p
9	27	100.0	206	1 KADA_MRTGA	Q97443 streptomyce
10	27	100.0	211	1 HISS_CORGL	Q06904 corynebacte
11	27	100.0	222	1 HISS_SPCRM	P16249 streptomyce
12	27	100.0	222	1 HISS_SPCRM	P16249 streptomyce
13	27	100.0	224	1 YMG3_CHEHL	Q26709 caenorhabdi
14	27	100.0	224	1 YMG3_CHEHL	Q26709 caenorhabdi
15	27	100.0	276	1 YMG3_CHEHL	Q26709 caenorhabdi
16	27	100.0	281	1 YMG3_CHEHL	Q26709 caenorhabdi
17	27	100.0	283	1 PHAB_PSEDL	P26495 pseudomonas
18	27	100.0	284	1 STAR_RAT	P97896 rattus norv
19	27	100.0	294	1 YQXK_BACSU	P38423 bacillus su
20	27	100.0	324	1 CCS1_CHEHL	P12114 caenorhabdi
21	27	100.0	334	1 YB96_MTCU	P71777 mycobacteri
22	27	100.0	374	1 OMFP_SPRHA	Q33980 serratia ma
23	27	100.0	393	1 EX33_MTCU	Q10699 mycobacteri
24	27	100.0	402	1 YMG3_CHEHL	Q26709 caenorhabdi
25	27	100.0	402	1 YMG3_CHEHL	Q26709 caenorhabdi
26	27	100.0	478	1 GSUB_AEATH	P46416 arabidopsia
27	27	100.0	479	1 CATP_PSEPU	Q59714 garchocera
28	27	100.0	482	1 CATP_PSEPU	Q59714 garchocera
29	27	100.0	488	1 U2AF_CHEHL	P90727 caenorhabdi
30	27	100.0	496	1 U2AF_CHEHL	P90727 caenorhabdi
31	27	100.0	510	1 CP46_RABIT	P14580 oryctolagus
32	27	100.0	511	1 CP47_RABIT	P14581 oryctolagus
33	27	100.0	519	1 CBX2_MOUSE	P30658 mus muscula

ALIGNMENTS

RESULT 1

ID	YMG3_SPCRM	STANDARD	PRT	93 AA
AC	Q05072	1994 (Rel. 28, Created)		
DT	01-FEB-1994	(Rel. 28, Last sequence update)		
DI	15-DEC-1998	(Rel. 37, Last annotation update)		
DE	HYPOTHETICAL PROTEIN IN MUTB 3' REGION (ORF-C) (FRAGMENT)			
OS	Streptomyces cinnamonensis			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae			
OX	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces			
NCBI_TaxID	19000			
PP	SEQUENCE FROM N.A.			
RA	Birch A. Leiser A.; Robinson J.A.			
RT	"Cloning, sequencing, and expression of the gene encoding			
RL	J. Bacteriol. 175:3511-3519(1993)			
CC	-/- SIMILARITY: BELONGS TO THE ARGX FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation			
CC	at the European Bioinformatics Institute, Cambridge, UK. The content is			
CC	not to be distributed outside the EMBL Outstation without the written			
CC	consent of the EMBL Outstation. The content is long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announc/			
CC	or send an email to license@sib-sib.ch)			
DR	EMBL; L10064; AAA03042.1			
DR	PIR; C40595; C40595			
FT	Hypothetical protein			
FT	NON-TER 93			
SQ	SEQUENCE 93 AA: 9838 MW; E667D064583845D3 CMO64;			

Query Match 100.0%; Score 27; DB 1; Length 93;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 VPGVG 5

Db 41 VPGVG 45

RESULT 2

ID	COXE_CHEHL	STANDARD	PRT	128 AA
AC	Q20779	1997 (Rel. 35, Created)		
DT	01-NOV-1997	(Rel. 35, Last sequence update)		
DI	15-JUL-1999	(Rel. 38, Last annotation update)		
DE	PROBABLE CYTOCHROME C OXIDASE POLYPEPTIDE VIA PRECURSOR (EC 1.9.3.1)			
OS	Caenorhabditis elegans			

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Peloderae; Caenorhabditis.
 CC NCBI_TaxID=6233;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN-BRISTOL N2;
 CC Submitted 11-NOV-1994 to the EMBL/GenBank/DBJ databases.
 CC Bentley D.;
 CC -1- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
 CC CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
 CC MITOCHONDRIAL ELECTRON TRANSPORT (BY SIMILARITY).
 CC 4 PROTONIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
 CC 4 PROTONS.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIA FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U12966; AAA20614.1;
 CC GenBank: F5408.2; CE01108.
 CC InterPro: IP0011349;
 CC Pfam: PF02046; COX6A; 1.
 CC PROSITE: PS01329; COX6A; 1.
 CC OXIDOREDUCTASE; Inner membrane; Mitochondrion; Transit peptide.
 CC TRANSIT ? MITOCHONDRION (POTENTIAL).
 CC CHAIN ? 128 PROBABLY CYTOCHROME C OXIDASE POLYPEPTIDE
 CC VTA.
 CC SEQUENCE 128 AA; 14743 MW; EF4E56A1CE56A233 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 128;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 116 VPGVG 120

RESULT 3
 ID Y5LH0RUY STANDARD; PRT; 181 AA.
 AC Q48509;
 DT 13-DEC-1998 (Rel. 37, Created).
 DE 15-DEC-1998 (Rel. 37, Last sequence update).
 DE 15-JUL-1999 (Rel. 38, Last annotation update).
 DE YCF65-LIKE PROTEIN PRECURSOR.
 OS Hordeum vulgare (Barley).
 CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae;
 CC Hordeum.
 CC NCBI_TaxID=4513;
 CC [1] SEQUENCE FROM N.A.
 CC STRAIN-CV; HANZA; TISSUE-Leaf;
 CC Bass W.R., Golz R., Boernert J.;
 CC "Analysis of randomly selected cDNAs reveals the expression of stress-
 CC and defence-related genes in the barley mutant albobrians";
 CC Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE YCF65 FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC Modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AJ222779; CAA10984.1;
 CC Hypothetical protein; Chloroplast; Transit peptide.
 CC TRANSIT ? CHLOROPLAST (POTENTIAL).
 CC CHAIN ? 181 YCF65-LIKE PROTEIN.
 CC SEQUENCE 181 AA; 19865 MW; B02DAC3792F7E8B5 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 181;
 Best Local Similarity 100.0%; Pred. No. 94;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
 DB 120 VPGVG 124

RESULT 4
 ID KADA_METIG STANDARD; PRT; 192 AA.
 AC P43408;
 DT 01-NOV-1995 (Rel. 32, Created).
 DT 01-CCF-1996 (Rel. 34, Last sequence update).
 DE 15-DEC-1998 (Rel. 37, Last annotation update).
 DE ADENYLATE KINASE (EC 2.7.4.3) (AIP-AMP TRANSPHOSPHORYLASE).
 GN ADKA_O3_AOK
 OS Methanococcus igneus.
 CC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus.
 CC NCBI_TaxID=2189;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE:9720879; PubMed=9055821;
 CC "Serber D.M., Haney P.J., Berk H., Lynn D., Konisky J.;
 CC "The complete kinase genes of M. voltae, M. thermolithotrophicus, M.
 CC "Methanococcus igneus and M. marisnigri define a new family of adenylate kinases";
 CC Gene 185:239-244(1997).
 CC [2]
 CC SEQUENCE OF 1-30.
 CC MEDLINE-95286473; PubMed=7768791;
 CC Rusnak P., Haney P., Konisky J.;
 CC "The adenylate kinases from a mesophilic and three thermophilic
 CC methanogenic members of the Archaea";
 CC J. Bacteriol. 177:2977-2981(1995).
 CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
 CC -1- SUBUNIT: MONOMER (PROBABLY).
 CC -1- TISSUE: CYTOPLASM; CYTOSOL.
 CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 70 TO 90 DEGREES
 CC CELSIUS.
 CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: U19881; AAC44862.1;
 CC GenBank: U19881; Kinase; ATP-binding
 CC NP BIND 10 18 ATP (BY SIMILARITY).
 CC CONFLICT 8 8 V->I (IN REF. 2).
 CC CONFLICT 20 20 T->I (IN REF. 2).
 CC SEQUENCE 192 AA; 21400 MW; AB2E3C9DB905E75 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 192;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 11 VPGVG 15

```

RESULT 5
KADA_METJA
ID KADA_METJA STANDARD; PRT; 192 AA.
AC P43409;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN ADXA OR ADX.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
NCBI_TaxID=2190;
[1]
[1]
RX MEDLINE-9720879; PubMed=9055921;
RA Fether D.H., Haney P.J., Beck H., Lynn D., Konisky J.;
RT "The adenylyate kinase genes of M. voltae, M. thermolithotrophicus, M.
jannaschii, and M. igneus define a new family of adenylyate kinases."
RL Gene 185:239-244(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE-95286473; PubMed=7768791;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.T., Weissbrock A.G., Merrick J.H., Glodek A.,
Piterback T.R., Klenz J.S., Peterson J.D., Sadow P.W., Adams K.,
Cotton M.D., Roberts K.M., Rust H.A., Xiang B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Wease C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii."
RL Science 273:1058-1073(1996).
[3]
RP SEQUENCE OF 1-30.
RX MEDLINE-95286473; PubMed=7768791;
RA Rusnak P., Haney P., Konisky J.;
RT "The adenylyate kinases from a mesophilic and three thermophilic
methanogenic members of the Archaea."
RL J. Bacteriol. 177:2977-2981(1995).
[4]
RX MEDLINE-177-2977-2981(1995).
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 70 TO 90 DEGREES
CELSIUS.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use. The European Bioinformatics Institute and the EMBL outstation
at the European Bioinformatics Institute have agreed to make this entry
available for commercial use. The copyright is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: U39882; AAC4864.1;
DR EMBL: U67499; AAB98470.1; AUT-INIT.
TX TIGR: XM0479;
KW Transferase; Kinase; ATP-binding.
FT CONFLICT 16 16 S->G (IN REF. 3).
SQ SEQUENCE 192 AA; 21772 MW; ECD533NDM4C8D599E CRC64;

Query Match 100.0%; Score 27; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 11 VPGVG 15

```

RESULT 6
KADA_METTL
ID KADA_METTL STANDARD; PRT; 192 AA.
AC P43410;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN ADXA OR ADX.
OS Methanococcus thermolithotrophicus.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
NCBI_TaxID=2186;
[1]
[1]
RX MEDLINE-9720879; PubMed=9055921;
RA Fether D.H., Haney P.J., Beck H., Lynn D., Konisky J.;
RT "The adenylyate kinase genes of M. voltae, M. thermolithotrophicus, M.
jannaschii, and M. igneus define a new family of adenylyate kinases."
RL Gene 185:239-244(1997).
[2]
RP SEQUENCE OF 1-20.
RX MEDLINE-95286473; PubMed=7768791;
RA Rusnak P., Haney P., Konisky J.;
RT "The adenylyate kinases from a mesophilic and three thermophilic
methanogenic members of the Archaea."
RL J. Bacteriol. 177:2977-2981(1995).
[4]
RX MEDLINE-177-2977-2981(1995).
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 60 TO 80 DEGREES
CELSIUS.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use. The European Bioinformatics Institute and the EMBL outstation
at the European Bioinformatics Institute have agreed to make this entry
available for commercial use. The copyright is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: U39880; AAC4864.1;
DR EMBL: U67499; AAB98470.1; AUT-INIT.
TX TIGR: XM0479;
KW Transferase; Kinase; ATP-binding.
FT NP-BIND 10 18 ATP (BY SIMILARITY).
SQ SEQUENCE 192 AA; 21461 MW; 7223378A3320B1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
|||||
DB 11 VPGVG 15

```

RESULT 7
KADA_METVO
ID KADA_METVO STANDARD; PRT; 192 AA.
AC P43411;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN ADXA OR ADX.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
NCBI_TaxID=2190;
[1]
[1]
RX MEDLINE-9720879; PubMed=9055921;
RA Fether D.H., Haney P.J., Beck H., Lynn D., Konisky J.;
RT "The adenylyate kinase genes of M. voltae, M. thermolithotrophicus, M.
jannaschii, and M. igneus define a new family of adenylyate kinases."
RL Gene 185:239-244(1997).
[2]
RP SEQUENCE OF 1-20.
RX MEDLINE-95286473; PubMed=7768791;
RA Rusnak P., Haney P., Konisky J.;
RT "The adenylyate kinases from a mesophilic and three thermophilic
methanogenic members of the Archaea."
RL J. Bacteriol. 177:2977-2981(1995).
[4]
RX MEDLINE-177-2977-2981(1995).
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 70 TO 90 DEGREES
CELSIUS.
CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL ADENYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on
its use. The European Bioinformatics Institute and the EMBL outstation
at the European Bioinformatics Institute have agreed to make this entry
available for commercial use. The copyright is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL: U39882; AAC4863.1;
DR EMBL: U67499; AAB98470.1; AUT-INIT.
TX TIGR: XM0479;
KW Transferase; Kinase; ATP-binding.
FT CONFLICT 16 16 S->G (IN REF. 3).
SQ SEQUENCE 192 AA; 21772 MW; ECD533NDM4C8D599E CRC64;

Query Match 100.0%; Score 27; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

GN ADKA OR ADK.
OS Methanococcus voltae.
CC Archaea: Crenarchaeota; Methanococcales; Methanococcaceae;
CC NCBI_TaxID=2188;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PS;
RA MEDLINE=97208879; PubMed=9055821;
RA Ferber D.M., Haney P.J., Berk H., Lynne D., Konisky J.;
RT "The adenylate kinase genes of M. voltae, M. thermolithotrophicus, M.
RT Janaschki, and M. igneus define a new family of adenylate kinases";
RL Gene 185:239-244(1997).
RN [1]
RP SEQUENCE OP 1-38
RA MEDLINE=95284573; PubMed=7768791;
RA Rusnak P., Haney P., Konisky J.;
RT "The adenylate kinases from a mesophilic and three thermophilic
RT methanogenic members of the Archaea";
RL J. Bacteriol. 177:2977-2981(1995).
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBUNIT: MONOMER (PROBABLE).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- MISCELLANEOUS: ACTIVE IN THE TEMPERATURE RANGE OF 30 TO 40 DEGREES
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ADENYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: U39879; AAC14855.1;
CC RefSeq: K11987; Kinase; ATP-binding.
CC NCBI_TaxID=31 31 G -> GG (IN REF. 2).
CC CONFLICT 31 31 G -> GG (IN REF. 2).
CC SEQUENCE 192 AA: 21303 MW: 95248CD178BD6A8E CRC64;
Query Match 100.0%; Score 27; DB 1; Length 192;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
Db 11 VPGVG 15
RESULT 8
ID KADA_AERPE STANDARD: PRT; 204 AA.
AC Q9YDD2;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE ADENYLATE KINASE (EC 2.7.4.3) (ATP-AMP TRANSPHOSPHORYLASE).
GN ADKA OR APEC0991.
OS Aeropyrum pernix.
CC Archaea: Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
CC NCBI_TaxID=55636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RA Kavarayasi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-NO K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamizawa M., Masuda S., Funahashi T., Tanaka T., Kudo H.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,

Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1";
CC -1- CATALYTIC ACTIVITY: ATP + AMP = ADP + ADP.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE ARCHAEL ADENYLATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: AP000060; BA079965.1;
CC RefSeq: K11987; Kinase; ATP-binding.
CC NCBI_TaxID=14 22
CC SEQUENCE 204 AA: 22175 MW: 279E5F8B081D13 CRC64;
Query Match 100.0%; Score 27; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPGVG 5
Db 15 VPGVG 19
RESULT 9
ID HISS_MYCLE STANDARD: PRT; 206 AA.
AC Q9X7C0;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE AMIDOTRANSFERASE HISH (EC 2.4.2.-).
GN HISH OR MCB1610.21.
OS Mycobacterium leprae.
CC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC NCBI_TaxID=17659;
RN [1]
RP SEQUENCE FROM N.A.
RA Seeger K.J., Harris D., James K.D., Parkhill J., Barrell B.G.,
RA Rajadream N.A.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE AMIDOTRANSFER FROM THE GLUTAMATES
CC INDUOXE GLYCEROL PHOSPHATE AND STAMINOLACTON- CARBOXYAMIDE
CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.
CC -1- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HISH FAMILY.
CC -1- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: AL049913; CAB43169.1;
CC InterPro: IPR000591;
CC Pfam: PF00117; GATase; 1.
CC PROSITE: PS00442; GATASE_TYPE 1;
CC ACT_SITE 83 83 BY SIMILARITY.
CC ACT_SITE 187 187 BY SIMILARITY.
CC ACT_SITE 189 189 BY SIMILARITY.

SQ SEQUENCE 206 AA; 21652 MW; 72D5994084F81536 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 206;

Best Local Similarity 100.0%; Pred. No. 1.1e+02; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

|||||

DB 47 VPGVG 51

RESULT 10

HIS_MYCTU

ID HIS_MYCTU STANDARD; PRT; 206 AA.

AC 01-OCT-2000 (Rel. 40, Created)

DT 01-OCT-2000 (Rel. 40, Last sequence update)

DE AMIDOTRANSFERASE HISH (EC 2.4.2.-)

GN HISH OR RV1602 OR MYC336.02C

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=H395987; PubMed=954320;

RA Cole S.T., Prosch R., Parbhall J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekle A.F.,

RA Badcock K., Haslam D., Brown D., Chillingworth T., Connor R.,

RA Davies R., de Lencastre A., Feltwell T., Genies S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagals K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.B.,

RA Taylor K., Whitehead S., Barrett B.G.;

RA "Deciphering the biology of Mycobacterium tuberculosis from the

RA complete genome sequence."

RA Nature 403:520-526 (2000).

RA CC -!- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES

CC IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE

CC RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.

CC -!- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.

CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).

CC -!- SIMILARITY: BELONGS TO THE HISH FAMILY.

CC -!- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; Z95586; CAB09092.1;

CC TubercuList; RV1602;

CC InterPro; IPR000981;

CC Pfam; PF00117; GATase; 1.

CC PROSITE; PS00442; GATASE_TYPE_1;

CC Histidine biosynthesis; Transferase; Glutamine amidotransferase.

CC ACT_SITE 87 187 BY SIMILARITY.

CC ACT_SITE 187 187 BY SIMILARITY.

CC ACT_SITE 189 189 BY SIMILARITY.

CC SEQUENCE 206 AA; 21418 MW; 4B2AF5C51BEE1447 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 206;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

|||||

DB 47 VPGVG 51

RESULT 12

HIS_STRCO

ID HIS_STRCO STANDARD; PRT; 222 AA.

AC P12249.1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DE AMIDOTRANSFERASE HISH (EC 2.4.2.-)

GN HISH OR SC4G6.20C

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

CC Actinomycetales; Streptomycineae; Streptomyces; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RA STRAIN=ATCC 13059 / AS019;

RA Jun S.I., Han M.S., Park Y.J., Lee S.K., Lee M.S.;

RA Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RA -!- FUNCTION: CATALYZES AN AMIDOTRANSFERASE REACTION THAT GENERATES

RA IMIDAZOLE-GLYCEROL PHOSPHATE AND 5-AMINOIMIDAZOL-4-CARBOXAMIDE

RA RIBONUCLEOTIDE, WHICH IS USED FOR PURINE SYNTHESIS.

RA -!- PATHWAY: FIFTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.

RA -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).

RA -!- SIMILARITY: CONTAINS 1 TYPE-1 GLUTAMINE AMIDOTRANSFERASE DOMAIN.

RA This SWISS-PROT entry is copyright. It is produced through a collaboration

RA between the Swiss Institute of Bioinformatics and the EMBL outstation -

RA the European Bioinformatics Institute. There are no restrictions on its

RA use by non-profit institutions as long as its content is in no way

RA modified and this statement is not removed. Usage by and for commercial

RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

RA or send an email to license@isb-sib.ch).

RA -----

RA EMBL; AF060586; AAC15231.1;

RA Pfam; PF00117; GATase; 1.

RA PROSITE; PS00442; GATASE_TYPE_1;

RA Histidine biosynthesis; Transferase; Glutamine amidotransferase.

RA ACT_SITE 82 82 BY SIMILARITY.

RA ACT_SITE 190 190 BY SIMILARITY.

RA ACT_SITE 192 192 BY SIMILARITY.

RA SEQUENCE 211 AA; 23139 MW; A935FEC4C2B49A1 CRC64;

Query Match 100.0%; Score 27; DB 1; Length 211;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VPGVG 5

|||||

DB 47 VPGVG 51

Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrall B.G.: "Deciphering the Biology of Mycobacterium tuberculosis from the Complete Genome Sequence". Nature 393:537-544(1998).

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>) or send an email to license@isb-sib.ch.

EMBL Z70692; C8A94667.1;
Trubercululin; RV2237;
Hypo-thetical protein; Transmembrane.
TRANSMEM 2 22 POTENTIAL.
TRANSMEM 168 188 POTENTIAL.
SEQUENCE 255 AA; 29067 MW; BOF422277C6B46A6 CRG64;

Query Match
Best Local Similarity 100.0%; Score 27; DB 1; Length 255;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGVG 5
DB 14 VPGVG 18

RESULT 14
DCOP_MYTCTU STANDARD; PRT: 274 AA.

ID DCOP_MYTCTU
AC P77898; A2610;
DT 01-NOV-1995 (Rel. 32; Created)
DI 01-NOV-1995 (Rel. 35; Last sequence update)
DE 30-MAY-2000 (Rel. 39; Last annotation update)
DE OXIDOLINE 5'-PHOSPHATE DECARBOXYLASE (EC 4.1.1.23) [OMP
PTPase] (OMPACTSE).
DE PROKARYOTIC ORNITHINYL D-CARBOXYPEPTIDASE
MYCOBACTERIUM TUBERCULOSIS H37Rv.
OS Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_Taxid=1773, 1765;
[1]
SEQUENCE FROM N.A.
RP SPECIES=M.tuberculosis; STRAIN=H37RV;
RC MEDLINE=98295987; PUBMED=9634230;
FX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekela F., Badcock N., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Felwell A., Freitag J., Gentles S., Hamlin N., Holroyd S., Hornsby A., Jagels K., Kelly T., McLean S., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrall B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence." Nature 393:537-544(1998).
[2]
SEQUENCE FROM N.A.
RP SPECIES=M.bovis; STRAIN=BCG;
RC MEDLINE=94042902; PUBMED=8226675;
FX Aloudini A., Huxson R.N., Young R.A.;
RA The ura locus and homologous recombination in Mycobacterium bovis BCG.;
RL Bacterial 175:782-789(1993).
RT -- PATHWAY ACTIVITY. OXIDOLINE 5'-PHOSPHATE = UMP + CO(2).
RT -- PATHWAY: SIXTH AND LAST STEP IN THE BIOSYNTHESIS OF PYRIDINES.
CC -- SIMILARITY. BELONGS TO THE OMP DECARBOXYLASE FAMILY.

QY 1 VPGVG 5
221 VPGVG 22

Search completed: April 24, 2001, 16:42:50
Job time: 456 sec

```

Query Match      100.0%; Score 27; DB 1; Length 274;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches      5; Conservative 0; Mismatches 0; Indels

```

Qy 1 VPGVG 5
|||
pb 215 VPGVG 219

RESULT 15	DOOP_MYCSM	STANDARD:	PRT: 276 AA.
ID	DOOP_MYCSM		
DOOP_MYCSM			
088323:			
30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)		
DT	01-OCT-2000 (Rel. 40, Last annotation update)		
DT	OROTIDINE 5'-PHOSPHATE DECARBOXYLASE (EC 4.1.1.23) (OMP		
DE	DECARBOXYLASE) (OMPDCASE).		
DE	MYCOBACTERIUM SMAGMATIS.		
DE	MYCOBACTERIUM SMAGMATIS.		
OC	Bacteria; Firmicutes; Actinobacteres; Actinobacteridae;		
OC	Actinomycetales; Corynebacteriaceae; Mycobacterium.		
OC	NCBI_TaxID=1772;		
OR	[1]		
OR	SEQUENCE FROM N.A.		
OR	STRAIN=ATCC 700084; / MC(2)155.		
OR	MDLING=971312763; PubMed=9169504.		
OR	Knipfer N., Seth A., Strider T.;		
OR	smagmatis 916 precise replacement of the pyrF gene.";		
OR	Plasmid 37.129-140(1997).		
RL	!! CATALYTIC ACTIVITY: OROTIDINE-5'-PHOSPHATE + H ₂ O + CO(2).		
CC	!! PATHWAY: SIXTH AND LAST STEP IN THE BIOSYNTHESIS OF PYRIMIDINES.		
CC	!! SUBUNIT: HOMODIMER (BY SIMILARITY).		
CC	!! SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outposts of		
CC	the European Bioinformatics Institute. There are no restrictions on		
CC	access to the data or to the software, but the data and for commercial		
CC	modification and this statement is not removed usage by		
CC	entities requires a license agreement (See http://www.isb-sib.ch/ann		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
CC	EMBL; U01572; RAH500157.1; "		
DR	InterPro; IPRO01754; "		
DR	Fam; PF00215; OMPdcase; 1.		
DR	PROSITE; PS00136; OMPDCASE; 1. Decarboxylase.		
DR	Protein name: Orotidine 5'-phosphate decarboxylase; EC: 4.1.1.23		
DR	Accession: U01572		
DR	Accession: 95		
DR	SEQUENCE 276 AA; 27815 MW; 41437843A32A896 CRC64;		
DR	50		

Query Match 100.0%; Score 27; DB 1; Length 276;

```

RESULT 2
US-07-609-716-1
; Sequence 1: Application US/07609716
; Patent No. 593385
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; TITLE OF INVENTION: Synthetic Protein Polymer
; NUMBER OF SEQUENCES: 118
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fehr, Honbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07609716
; FILING DATE: 06-NOV-1990
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REFERENCE/DOCKET NUMBER: 20015
; REGISTRATION NUMBER: A-55186-3/BIR
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-07-609-716-1

Query Match 100.0%; Score 23; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
Db 1 VPGG 4

RESULT 3
US-08-483-236-18
; Sequence 18, Application US/08483236
; Patent No. 593385
; GENERAL INFORMATION:
; APPLICANT: Labroo, Virender
; APPLICANT: Busby, Sharon
; TITLE OF INVENTION: Transglutaminase Cross-Linkable
; TITLE OF INVENTION: Polypeptides and Methods Relating Thereto
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08483236
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-09c1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6673
; TELEFAX: 206-442-6678
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-483-236-18

Query Match 100.0%; Score 23; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
Db 1 VPGG 4

RESULT 4
US-08-911-364-7
; Sequence 7, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION:
; APPLICANT: Rothstein, Steven J
; APPLICANT: Kelly, Fred W
; APPLICANT: Rothstein, Steven J
; TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911.364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,552
; FILING DATE: 07-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 041082/0104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids

```

```

; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-911-364.7

```

```

Query Match      100.0%; Score 23; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VPGG 4
    ||||
DB 1 VPGG 4

```

RESULT 5

```

US-08-735-692-9
; Sequence 9, Application US/08735692B
; Patent No. 5972406
; GENERAL INFORMATION:
; APPLICANT: Urry, Dan W.
; APPLICANT: Sherry, Peter R.
; APPLICANT: Prasad, Kari
; TITLE OF INVENTION: Bioclastomers Suitable as Food Product Additives
; FILE REFERENCE: BERL-016/01US
; CURRENT FILING DATE: 1995-10-16
; EARLIER APPLICATION NUMBER: 08/423,517
; EARLIER FILING DATE: 1995-04-14
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-08-735-692-9

```

```

Query Match      100.0%; Score 23; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 VPGG 4
    ||||
1 VPGG 4

```

RESULT 6

```

US-08-542-051-3
; Sequence 3, Application US/08542051F
; Patent No. 6004782
; GENERAL INFORMATION:
; APPLICANT: Daniell, Henry
; APPLICANT: McPherson, David T.
; APPLICANT: Urry, Dan W.
; APPLICANT: Xu, Jie
; TITLE OF INVENTION: Hyperexpression of Bioelastic Polypeptides
; FILE REFERENCE: BERL-016/01US
; CURRENT APPLICATION NUMBER: US/08/542,051F
; CURRENT FILING DATE: 1995-10-13
; EARLIER APPLICATION NUMBER: 08/423,642
; EARLIER FILING DATE: 1995-04-14
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:synthetic

```

US-08-542-051-3

```

Query Match      100.0%; Score 23; DB 3; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VPGG 4
    ||||
DB 1 VPGG 4

```

RESULT 7

```

US-08-963-168C-29
; Sequence 29, Application US/08963168C
; Patent No. 6127166
; GENERAL INFORMATION:
; APPLICANT: Raley, Hagan
; APPLICANT: Cui, Yujun
; APPLICANT: Wang, Yujun
; TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/963,168C
; FILING DATE: 03-NOV-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Fasse, Peter J.
; REGISTRATION NUMBER: 32,983
; ELECTRONIC FILING DATE: 07/91/059001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-5906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-963-168C-29

```

```

Query Match      100.0%; Score 23; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 VPGG 4
    ||||
DB 1 VPGG 4

```

RESULT 8

```

US-08-475-411A-1
; Sequence 1, Application US/08475411A
; Patent No. 610072
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; APPLICANT: Cappello, Joseph
; TITLE OF INVENTION: Functional Recombinantly Prepared
; OTHER INFORMATION: Synthetic Protein Polymer

```

```

; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,411A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/269,429
; FILING DATE: 09-NOV-1988
; APPLICATION NUMBER: US 06/927,258
; FILING DATE: 04-NOV-1986
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-9/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-398-3249
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 4 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-475-411A-1

Query Match 100.0%; Score 23; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFGG 4
DB 1 VFGG 4

; SEQUENCE INFORMATION:
; SEQUENCE 1, Application US/08/478029A
; Patent No. 6184348
; GENERAL INFORMATION:
; APPLICANT: Ferrari, Franco A.
; TITLE OF INVENTION: Functional Recombinantly Prepared
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,029A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/609,716
; FILING DATE: 06-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/259,429
; FILING DATE: 06-NOV-1988
; APPLICATION NUMBER: US 07/114,618
; FILING DATE: 29-OCT-1987
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-55186-8/RFT/MTK
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-781-3249
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 4 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-478-029A-1

Query Match 100.0%; Score 23; DB 4; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFGG 4
DB 1 VFGG 4

; SEQUENCE INFORMATION:
; SEQUENCE 1, Application US/07/184,407
; Patent No. 5250516
; APPLICANT: URRY, DAN W.
; TITLE OF INVENTION: BIOLASTOMERIC MATERIALS SUITABLE FOR
; THE PROTECTION OF BURN AREAS OR THE PROTECTION OF WOUND
; REPAIRS FROM EXCESSIVE LOSS OF FLUID AND OCCURRENCE OF ADHESIONS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/184,407
; FILING DATE: 21-APR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 900,895
; FILING DATE: 27-AUG-1986
; APPLICATION NUMBER: 853,212
; FILING DATE: 17-APR-1986
; SEQ ID NO: 3:
; LENGTH: 4
; 5250516-3

Query Match 100.0%; Score 23; DB 6; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFGG 4
DB 1 VFGG 4

```

Db 1 VPEG 4

RESULT 11
US-08-340-283-182
; Sequence 182, Application US/08340283
; Patent No. 5861318
; GENERAL INFORMATION:
; APPLICANT: Elhammer, Ake P.
; TITLE OF INVENTION: A SCINTILLATION PROXIMITY ASSAY FOR
; TITLE OF INVENTION: N-ACETYL GALACTOSAMINYLTRANSFERASE ACTIVITY
; NUMBER OF SEQUENCES: 205
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia and Upjohn, Inc., Intellect. Prop. Law
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,283
; FILING DATE: 436
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Wootton, Thomas A.
; REGISTRATION NUMBER: 35,004
; REFERENCE/DOCKET NUMBER: 4828
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (616) 385-7914
; TELEFAX: (616) 385-6897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 182:
; LENGTH: 7 amino acids
; TYPE: amino acid
; TOPOLOGY: single
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; US-08-340-283-182

Query Match 100.0%; Score 23; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPEG 4
Db 4 VPEG 7

RESULT 12
US-09-147-933-29
; Sequence 29, Application US/09147933A
; Patent No. 6168917
; GENERAL INFORMATION:
; APPLICANT: Kilpatrick, David
; TITLE OF INVENTION: DETECTION AND IDENTIFICATION OF
; TITLE OF INVENTION: DETECTION OF OLIGO NUCLEOTIDES
; FILE REFERENCE: 6242/US
; CURRENT APPLICATION NUMBER: US/09/147,933A
; CURRENT FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: PCT/US97/17734
; EARLIER FILING DATE: 1997-10-01
; EARLIER APPLICATION NUMBER: U. S. 60/027,353

; EARLIER FILING DATE: 1996-10-02
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 7
; TYPE: PRT
; ORIGIN: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide
US-09-147-933-29

Query Match 100.0%; Score 23; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPEG 4
Db 3 VPEG 6

RESULT 13
US-08-340-283-64
; Sequence 64, Application US/08340283
; Patent No. 5861318
; GENERAL INFORMATION:
; APPLICANT: Elhammer, Ake P.
; TITLE OF INVENTION: A SCINTILLATION PROXIMITY ASSAY FOR
; TITLE OF INVENTION: N-ACETYL GALACTOSAMINYLTRANSFERASE ACTIVITY
; NUMBER OF SEQUENCES: 205
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pharmacia and Upjohn, Inc., Intellect. Prop. Law
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: U.S.A.
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,283
; FILING DATE: 436
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Wootton, Thomas A.
; REGISTRATION NUMBER: 35,004
; REFERENCE/DOCKET NUMBER: 4828
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (616) 385-7914
; TELEFAX: (616) 385-6897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 64:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
US-08-340-283-64

Query Match 100.0%; Score 23; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPEG 4

Db 1 VPGG 4
||||
RESULT 14
US-08-282-758B-32
; Sequence 32, Application US/08282758B
; Patent No. 5670310
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; TITLE OF INVENTION: Methods and Compositions for
; TITLE OF INVENTION: Differential Diagnosis of Chronic and Acute Hepatitis C
; TITLE OF INVENTION: Virus Infection
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,758B
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0100
; TELEPHONE: (404) 818-3700
; TELEFAX: (404) 818-3799
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-282-758B-32
Query Match 100.0%; Score 23; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPGG 4
Db 7 VPGG 10
RESULT 15
US-08-282-758B-14
; Sequence 14, Application US/08282758B
; Patent No. 5670310
; GENERAL INFORMATION:
; APPLICANT: Fields, Howard A.
; TITLE OF INVENTION: Methods and Compositions for
; TITLE OF INVENTION: Differential Diagnosis of Chronic and Acute Hepatitis C
; TITLE OF INVENTION: Virus Infection
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jones & Askew
; STREET: 191 Peachtree Street, 37th Floor

; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.A.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,758B
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Jamie L.
; REGISTRATION NUMBER: 32,467
; REFERENCE/DOCKET NUMBER: 03063-0100
; TELEPHONE: (404) 818-3700
; TELEFAX: (404) 818-3799
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-282-758B-14
Query Match 100.0%; Score 23; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 VPGG 4
Db 9 VPGG 12
Search completed: April 24, 2001, 16:36:25
Job time: 305 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:42:03 ; Search time 74.56 seconds
(without alignments)
3.687 Million cell updates/sec

Title: US-09-340-736-7
Perfect score: 23
Sequence: 1 VPEG 4

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR-67.1
1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	23	100.0	7	2 B61491	seed protein ws-5
2	23	100.0	15	2 PC2215	fibrinogenolytic p
3	23	100.0	27	2 A28321	hypothetical prote
4	23	100.0	30	2 S14062	hypothetical prote
5	23	100.0	31	2 I52232	tau protein - huma
6	23	100.0	41	2 T36863	probable two-compo
7	23	100.0	45	2 S20197	r-complex protein
8	23	100.0	46	2 T34928	tracheal mucin gly
9	23	100.0	56	2 T34928	tracheal mucin gly
10	23	100.0	61	2 S03097	neuro protein
11	23	100.0	62	2 S37135	class II histocomp
12	23	100.0	63	2 T04493	hypothetical prote
13	23	100.0	69	2 F49410	t-complex polypept
14	23	100.0	70	2 T36407	hypothetical prote
15	23	100.0	81	2 B69958	hypothetical prote
16	23	100.0	83	2 S04884	seed protein (clon
17	23	100.0	84	2 T07939	embryonic abundant
18	23	100.0	91	2 S16249	embryonic abundant
19	23	100.0	91	2 T09293	embryonic abundant
20	23	100.0	91	2 T09293	hypothetical prote
21	23	100.0	91	2 T21508	hypothetical prote
22	23	100.0	91	2 T21508	Nut/pudix family
23	23	100.0	92	1 S33700	ribosomal protein
24	23	100.0	92	2 S23577	embryonic abundant
25	23	100.0	92	2 JQ2273	embryonic abundant
26	23	100.0	92	2 S34802	embryogenic abunda
27	23	100.0	92	2 S71372	embryonic abundant
28	23	100.0	93	2 S42333	embryonic abundant
29	23	100.0	93	2 S45303	g619.1b protein -

EMF5 protein - whe
embryonic abundant
Em protein - wheat
LSD ribosomal prot
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
embryonic abundant
conserved hypothet
late-embryogenesis
tubulin alpha chai
hypothetical prote
perchloric acid-so
embryonic abundant
hypothetical prote

ALIGNMENTS

RESULT 1
ES1491
seed protein ws-5 - winged bean (fragment)
C:Species: Esophocarpus tetragonolobus (winged bean)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-Oct-1994
C:Accession: ES1491
J.Hirano, H.
J. Protein Chem. 8, 115-130, 1989
A:Title: Microsequence analysis of winged bean seed proteins electrophoretically separated from two
A:Accession: ES1491
A:Accession: ES1491
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-7 <EIR>
C:Keywords: glycoprotein; seed

Query Match 100.0%; Score 23; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPEG 4
DB 3 VPEG 6
RESULT 2
PC2215
fibrinogenolytic proteinase A2 (EC 3.4.21.-) - western diamondback rattlesnake (fragm
N:Alternate names: alpha-fibrinogenase A2
C:Species: Crotalus atrox (western diamondback rattlesnake)
C:Date: Feb-1995
C:Accession: PC2215
C:Accession: PC2215
C:Keywords: hydrolase; serine proteinase
Biochem. Biophys. Res. Commun. 201, 1414-1423, 1994
A:Title: Isolation of multiple isoforms of alpha-fibrinogenase from the western diamo
A:Reference number: PC2214; MUID:94295418
A:Accession: PC2215
A:Molecule type: protein
A:Residues: 1-15 <KUN>
C:Keywords: hydrolase; serine proteinase

Query Match 100.0%; Score 23; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VPEG 4
DB 1 VPEG 4

RESULT 3
A28391
hypothetical protein 1 - Streptomyces albidoflavus (fragment)
C:Species: Streptomyces albidoflavus
C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 11-Jan-2000
C:Accession: A28391
R:Long, C.M.; Virolle, M.J.; Chang, S.Y.; Chang, S.; Bibb, M.J.
J. Bacteriol. 159, 5745-5754, 1987
A:Title: Alpha-amylose gene of Streptomyces limosus: nucleotide sequence, expression and
A:Reference number: A91856; MUID:86036792
A:Accession: A28391
A:Molecule type: DNA
A:Residues: 1-27 <LON>
A:Note: the source is designated as Streptomyces limosus
C:Superfamily: lac repressor

Query Match 100.0%; Score 23; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
DB 24 VFGG 27

RESULT 4
S14062
hypothetical protein 1 - Streptomyces griseus (fragment)
C:Species: Streptomyces griseus
C>Date: 19-Mar-1997 #sequence_revision 26-Feb-1998 #text_change 11-Jan-2000
C:Accession: S14062
R:Yigal, T.; Gil, J.A.; Daza, A.; Garcia-Gonzalez, M.D.; Martin, J.F.
Mol. Gen. Genet. 225, 278-286, 1991
A:Title: Cloning, characterization and expression of an alpha-amylose gene from Streptomyces
A:Reference number: S14062; MUID:91172128
A:Accession: S14062
A:Molecule type: DNA
A:Residues: 1-30 <VIG>
A:Experimental source: strain IMU3570
C:Superfamily: lac repressor

Query Match 100.0%; Score 23; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
DB 27 VFGG 30

RESULT 5
I52232
tau protein - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1995 #sequence_revision 02-Jul-1995 #text_change 13-Aug-1999
C:Accession: I52232
R:Mori, H.; Hamada, Y.; Kawaguchi, M.; Honda, T.; Kondo, J.; Ihara, Y.
Biochem. Biophys. Res. Commun. 159, 1221-1226, 1989
A:Title: A distinct form of tau is selectively incorporated into Alzheimer's paired helix
A:Reference number: I52232; MUID:8919714
A:Accession: I52232
A:Molecule type: protein
A:Residues: 1-31 <RES>
A:Cross-references: NID:G602470; PIDN:RAA57264.1; PID:G602471
C:Superfamily: microtubule-associated protein tau, MAP2/tau repeat homology

Query Match 100.0%; Score 23; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
DB 26 VFGG 29

RESULT 6
T36863
Probable two-component sensor kinase - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1998 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36863
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z21616
A:Accession: T36863
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-41 <SEE>
A:Cross-references: EMBL:AL079332; PIDN:CAB45294.1; GSPDB:GN00070; SCOPDB:SC15.10C
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOPDB:SC15.10C

Query Match 100.0%; Score 23; DB 2; Length 41;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
DB 38 VFGG 41

RESULT 7
S28397
complex protein 5 - bovine (fragments)
C:Species: Bos taurus (cattle)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Mar-1999
C:Accession: S28397
R:Frydman, J.; Nimmesgern, B.; Erdjument-Bromage, H.; Wall, J.S.; Tempst, P.; Hartl, F.
EMBO J. 11, 4767-4778, 1992
A:Title: Function in protein folding of TRIC, a cytosolic ring complex containing TCP
A:Reference number: S28395; MUID:93099850
A:Accession: S28397
A:Molecule type: protein
A:Residues: 1-20; 21-45 <FRY>
C:Superfamily: molecular chaperone t-complex-type
C:Keywords: molecular chaperone

Query Match 100.0%; Score 23; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
DB 29 VFGG 32

RESULT 8
A61384
tracheal mucin glycoprotein - pig (fragments)
C:Species: Sus scrofa domestica domestica (pig)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Sep-1994
C:Accession: A61384
R:Sapadala, S.; Kim, D.; Brewer, J.M.; Mendicino, J.
Mol. Cell. Biochem. 102, 71-93, 1991
A:Title: Subunit structure of deglycosylated human and swine trachea and Cowper's gla
A:Reference number: A61384; MUID:91270244
A:Accession: A61384
A:Status: preliminary

A:Molecule type: protein
A:Residues: 1-47 <SAS>
C:Keywords: glycoprotein

Query Match 100.0%; Score 23; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 23 VPGG 26

RESULT 9
T36928
Hypothetical protein SC17.32 - Streptomyces coelicolor
A:Species: Streptomyces coelicolor
A:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
A:Accession: T36928
A:Submitted to the EMBL Data Library, July 1999
A:Reference number: 221574
A:Accession: T36928
A:Status: Preliminary
A:Residues: 1-56 <SEP>
A:Cross-references: EMBL:AL096743; PIDN:CB46415.1; GSPDB:GN00070; SC0EDB:SC17.32
A:Experimental source: strain A3(2)
A:Genetics:
A:Gene: SC0EDB:SC17.32

Query Match 100.0%; Score 23; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 5 VPGG 8

RESULT 10
S03097
aerC protein - Aeromonas sobria
A:Species: Aeromonas sobria
A:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 08-Oct-1999
A:Accession: S03097
A:Submitted to the EMBL Data Library, July 1989
A:Reference number: S03097; MID:8901382
A:Accession: S03097
A:Molecule type: DNA
A:Residues: 1-61 <HUS>
A:Cross-references: EMBL:I00559; MID:q39016; PIDN:CAA68641.1; PID:q39017
A:Genetics:
A:Gene: aerC

Query Match 100.0%; Score 23; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 35 VPGG 38

RESULT 11
S37135
class II histocompatibility antigen DR beta-6 chain - rhesus macaque (fragment)
A:Species: Macaca mulatta (rhesus macaque)

C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C:Accession: S37135
R:Slenderdregt, B.L.; Otting, N.; van Besouw, N.; Jonker, M.; Bontrop, R.E.
submitted to the EMBL Data Library, September 1993

A:Description: Expansion of rhesus macaque DRB regions by duplication.

A:Reference number: S37135
A:Accession: S37135
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-62 <SLI>
A:Cross-references: EMBL:Z26183

Query Match 100.0%; Score 23; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 41 VPGG 44

RESULT 12
T04493
Hypothetical protein F8P16.80 - Arabidopsis thaliana
A:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
A:Accession: T04493
A:Submitted to the EMBL Data Library, April 1998
A:Reference number: Z15375
A:Accession: T04493
A:Molecule type: DNA
A:Residues: 1-63 <BEV>
A:Cross-references: EMBL:AL021633
A:Experimental source: cultivar Columbia; BAC clone F8P16
A:Genetics:
A:Map position: 4
A:Note: F8P16.80

Query Match 100.0%; Score 23; DB 2; Length 63;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 7 VPGG 10

RESULT 13
F49410
Complex polypeptide 1 homolog (peak 4 fraction) - rabbit (fragments)
A:Species: Oryctolagus cuniculus (white rabbit)
C:Date: 21-Sep-1994 #sequence_revision 18-Nov-1994 #text_change 04-Sep-1998
A:Accession: F49410
R:Romelaere, H.; Van Troys, M.; Gao, Y.; Melki, R.; Cowan, N.J.; Vandekerckhove, J.;
Proc. Natl. Acad. Sci. U.S.A. 90, 11975-11979, 1993
A:Title: Sukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and seven x
A:Reference number: A49410; MID:94089752
A:Accession: F49410
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-69 <KOE>
A:Experimental source: rabbit reticulocyte
A:Note: sequence modified after extraction from NCB1 backbone
A:Note: sequence extracted from NCB1 backbone (NCBP:141034, NCBIP:141036)
C:Superfamily: molecular chaperone t-complex-type

Query Match 100.0%; Score 23; DB 2; Length 69;
Best Local Similarity 100.0%; Pred. No. 3e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VFGG 4
Db 39 VFGG 42
|||||
RESULT 14
T36407
Hypothetical protein SCF34.08c - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1997 #sequence_revision 03-Dec-1999 #text_change 19-May-2000
E:Saunders D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1999
A:Reference number: 221506
A:Accession: T36407
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-70 <SAU>
A:Cross-references: EMBL:AL109974; PIDN:CA55319.1; GSPDB:GN00070; SCOPDB:SCF34.08c
Experimental source: strain A3(2)
C:Genetics:
C:Gene: SCOPDB:SCF34.08c
C:Superfamily: Escherichia coli hypothetical 8k protein (fas 3' region)

Query Match 100.0%; Score 23; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFGG 4
Db 32 VFGG 35
|||||
RESULT 15
B69958
Hypothetical protein yggY - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: B69958
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azavedo, V.; Berto
C.; Bron, S.; Brouillet, S.; Buschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galle
lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.
Roetter, P.; Rongberg, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapicque, A.; Lardinois
A:Authors: Lauber, O.; Lazarevic, A.; Lee, S.H.; Levine, A.; Liu, H.; Masuda, S.; Maue
A.; Ogata, N.; Ogata, A.; Oudeguy, B.; Park, S.H.; Parco, J.; Ohn, I.M.; Pettele
A:Authors: Rivolt, C.; Ribi, E.; Rocco, R.; Rocco, R.; Rocco, R.; Rocco, R.; Rocco
A:Authors: Schleich, S.; Schroter, R.; Scoffone, P.; Sekiguchi, J.; Sekovska, A.; Sord
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MID:98044033
A:Accession: B69958
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-81 <XUN>
A:Cross-references: GB:259116; GB:AL009126; NID:92634723; PIDN:CA314409.1; PID:92634912
Experimental source: strain 108
C:Genetics:
A:Gene: yggY

Query Match 100.0%; Score 23; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFGG 4

Db 75 VFGG 78
|||||

Search completed: April 24, 2001, 16:42:04
Job time: 469 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: April 24, 2001, 16:42:50 ; Search time 44.88 Seconds
(without alignments)
3.035 Million cell updates/sec

Title: US-09-340-736-7

Perfect score: 23

Sequence: 1 VPGG 4

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	23	100.0	23	1 NEUO_LITCE	P81872 litoria cae
2	23	100.0	27	1 GRP_CANFA	P08989 canis fami
3	23	100.0	29	1 TLP_ACTDE	P81370 actinidia d
4	23	100.0	61	1 AERC_AERSO	P09165 aeromonas s
5	23	100.0	81	1 PRAC_CENYA	P81060 penaeus van
6	23	100.0	81	1 PRAC_CENYA	P81060 penaeus van
7	23	100.0	83	1 SPPE_EPRSP	P11573 babianus su
8	23	100.0	89	1 PRSN_EPRSP	P33570 pseudomona
9	23	100.0	90	1 LSN5_HDVAN	O93479 homo sapien
10	23	100.0	91	1 LSN5_MATZE	P45517 zea mays (m
11	23	100.0	91	1 ENL_PICGL	Q40854 picea glauc
12	23	100.0	92	1 ENL_AARAH	Q02973 arabidopsis
13	23	100.0	92	1 ENL_DAOCA	P17635 daucus caro
14	23	100.0	92	1 LELO_HELAN	P46514 helianthus
15	23	100.0	92	1 RL4L_HALMA	P22411 halocaulula
16	23	100.0	93	1 ENL_WHEAT	P04568 triticum ae
17	23	100.0	93	1 ENL_WHEAT	P04568 triticum ae
18	23	100.0	93	1 ENL_WHEAT	P04568 triticum ae
19	23	100.0	93	1 L192_HORVU	Q05190 hordeum vul
20	23	100.0	93	1 L192_HORVU	P46532 hordeum vul
21	23	100.0	93	1 LSN5_YEAST	P40089 saccharomyc
22	23	100.0	93	1 RL4L_ARCFU	O28936 archaeglob
23	23	100.0	94	1 ENL_WHEAT	P22701 triticum ae
24	23	100.0	95	1 ENL_ORYSA	P46520 oryza sativ
25	23	100.0	102	1 LE12_GOSHI	P09443 gossypium h
26	23	100.0	102	1 MEN_ORYSA	P49030 oryza sativ
27	23	100.0	111	1 SC4_SORCO	P16934 schizopyll
28	23	100.0	112	1 ENL_ECOLI	P07709 escherichia
29	23	100.0	124	1 ENL_ECOLI	P07709 escherichia
30	23	100.0	124	1 ENL_ECOLI	P07709 escherichia
31	23	100.0	125	1 Y43N_SHISN	P50358 rhizobium s
32	23	100.0	126	1 Y43N_SHISN	P50358 rhizobium s
33	23	100.0	126	1 Y43N_SHISN	P50358 rhizobium s

P76297 escherichia
P40728 salmonella
P55499 rhizobium s
Q02400 hordeum vul
P52759 rattus norv
P57391 mus musculu
P51433 rattus norv
P50661 capra hircu
P57961 mus musculu
Q04001 psocoptera
P50439 caenorhabdi
Q07187 arabidopsis

ALIGNMENTS

RESULT 1
NEUO_LITCE 1
ID REFID LITCE STANDARD: PRT: 23 AA.
PC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NEUROMEDIN U-23 (NMU-23).
OS Litoria caerulea.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Litoria.
CX NCBI_TaxID=30344;
RN [1]
RP SwissProt.
RC NEUO_Skin.
RC MEDLINE=2013182; PubMed=10671478;
RA Salomon A.L., Johnson A.H., Bismart M., McMurray G., Nandha K.A.,
Bloom S.R., Shaw C.;
RT Isolation, structural characterization and bioactivity of a novel
neuroedin U analog from the defensive skin secretion of the
RT Australasian tree frog Litoria caerulea.;
RL J. Biol. Chem. 275:4549-4554(2000).
CC -/- FUNCTION: STIMULATE UTERINE SMOOTH MUSCLE CONTRACTION AND CAUSE
CC -/- SELECTIVE VASOCONSTRICTION.
CC -/- SIMILARITY: BELONGS TO THE NMU FAMILY.
DR InterPro: IPR001942.;
DR PROSITE: PS00367; NMU; 1.
FT NCD_RES 23
SQ SEQUENCE 23 AA: 2581 MW: A94958415CB58DC3 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 7 VPGG 10

RESULT 2
GRP_CANFA 2
ID GRP_CANFA STANDARD: PRT: 27 AA.
AC P08989;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE GASTRIN-RELEASING PEPTIDE (GRP) [CONTAINS: NEUROMEDIN C (GRP-10)].
OS Canis familiaris (Dog).
OC Chordata; Mammalia; Eutheria; Carnivora; Fissipedia; Canis.
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Fissipedia; Canis.
CX NCBI_TaxID=9615;
RN [1]

RP SEQUENCE
RA MEDLINE-43213518; PubMed-6553532;
RA Smith R. J., Walsh J. H., Chew P., Clark B., Hawke D.,
RA "Amino acid sequences of three bombesin-like peptides from canine
RA intestine extracts."
RA J. Biol. Chem. 258:5582-5588(1983).
CC -1- FUNCTION: GRP STIMULATES GASTRIN RELEASE AS WELL AS OTHER
CC GASTROINTESTINAL HORMONES.
CC -1- SIMILARITY: BELONGS TO THE BOMBESIN/NEUREGONIN B/RANATENSIN
CC FAMILY.
CC InterPro: IPR000874; -
CC Pfam: PF02044; Bombesin: 1.
CC PROSITE: PS00257; BOMBESIN; 1.
CC NeuroMedin C.
CC NCBI TaxID: 27
CC SEQUENCE 27 AA; 2689 MW; 9D531756157C7D65 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 VPGG 4
Db 3 VPGG 6

RESULT 3
ID TLP_ACTOR STANDARD; PRT: 29 AA.
AC P81370;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE THAUMATIN-LIKE PROTEIN (FRAGMENT).
OS Actinidia deliciosa (Kiwi).
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Magnoliophyta; Eudicotyledons; Core eudicots; Asteridae; Ericales;
CC Actinidia.
CC NCBI_TaxID:3627;
RN [1]
RP SEQUENCE.
RC STRAIN-CV. HAYWARD; TISSUE-Stem plug;
RA Wurms K.V., Greenwood D.R., Sharrock K.R., Long P.G.;
RA "Thaumatococcus-like protein in kiwifruit."
RL J. Sci. Food Agric. 79:1448-1452(1999).
CC -1- TISSUE SPECIFICITY: WOODY STEM PLUG.
CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
CC HSP: P25971; INUN.
CC InterPro: IPR00338; -
CC PROSITE: PS00314; THAUMATIN; PARTIAL.
DR PROSITE: PS00314; THAUMATIN; PARTIAL.
FT NON_TER 29
SQ SEQUENCE 29 AA; 3074 MW; B99F916F50AC377B CRC64;

Query Match 100.0%; Score 23; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 VPGG 4
Db 18 VPGG 21

RESULT 4
ID AERC_AERSO STANDARD; PRT: 61 AA.
AC P09165;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)

DT 01-MAR-1989 (Rel. 10, Last annotation update)
DE AEROLISIN REGULATORY PROTEIN.
OS Aeromonas sobria
CC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;
CC Aeromonas.
CC NCBI_TaxID=646;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AB3;
RA MEDLINE-89013892; PubMed-2459581;
RA Husslein V., Huhle B., Jarchau T., Lurz R., Gobel W., Chakraborty T.;
RA "Nucleotide sequence and transcriptional analysis of the aerChcA
RA region of Aeromonas sobria encoding aerolysin and its regulatory
RA region."
RA Mol. Microbiol. 2:507-517(1988).
CC -1- FUNCTION: REGULATION OF THE EXPRESSION OF AEROLYSIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC EMBL: J00559; CRA6841.1;
CC EXP: J03097; S03097;
CC SEQUENCE 61 AA; 6452 MW; DCF7C7DBA7ED752 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 VPGG 4
Db 35 VPGG 38

RESULT 5
ID PNC3_PENVA STANDARD; PRT: 81 AA.
AC P81060;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PENNAIDIN-3C PRECURSOR (P3-C).
OS Penaeus vannamei (Fenoid shrimp) (European white shrimp).
CC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
CC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
CC Penaeidae; Eus.
CC NCBI_TaxID=6689;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Hemocyte;
RA MEDLINE-98013209; PubMed-9353298;
RA Destombeux D., Bulet P., Joew D., van Dorsellaer A., Rodriguez J.,
RA Bachere E.;
RA "Penaeidins, a new family of antimicrobial peptides isolated from the
RA shrimp Penaeus vannamei (Decapoda)."
RL J. Biol. Chem. 272:28398-28406(1997).
CC -1- TISSUE SPECIFICITY: HEMOCYTES.
CC -1- TISSUE SPECIFICITY: HEMOCYTES.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL; Y14926; CAA75145.1; .
KW Antibiotic; Amidation; Signal.
FT SIGNAL 1 ? POTENTIAL.
FT PROPEP ? 19 BY SIMILARITY.
FT CHAIN 20 80 PENAETIDIN-3C.
FT MOD_RES 20 20 PYROLIDONE CARBOXYLIC ACID (BY
FT MOD_RES 80 80 SIMILARITY).
FT MOD_RES 80 80 SIMILARITY). (G-91 PROVIDE AMIDE GROUP) (BY
FT MOD_RES 80 80 SIMILARITY).
FT DOMAIN 29 46 PRO-RICH.
FT PRO-RES 29 46 PRO-RICH.
FT PRO-RES 29 46 PRO-RICH.
SQ SEQUENCE 81 AA; 8637 MW; 2950B36163B92C36 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VFGG 4
|||||
39 VFGG 42

RESULT 5
ID YQST_BACSU STANDARD; PRT; 81 AA.
AC P54502;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE HYPOTHETICAL 9.4 KDA PROTEIN IN SODA-COMGA INTERGENIC REGION.
OR Nucleotide subunit.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RA Kobayashi Y., Mizuno M., Masuda S., Takenaru K., Hosono S.,
RA Sato T., Takeuchi M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D94432; BAA12528.1; .
DR EMBL; Z95116; CAB14409.1; .
DR SUBLIST; B61692; YGGY.
KW Chemical protein; Transmembrane
FT TRANSMEM 54 74 POTENTIAL.
FT TRANSMEM 54 74 POTENTIAL.
SQ SEQUENCE 81 AA; 9422 MW; B22D7BF840DEA29 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VFGG 4
|||||
75 VFGG 78

RESULT 7
ID SEEP_RAPSA STANDARD; PRT; 83 AA.
AC P11573;
DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE LATE SEED MATURATION PROTEIN P886.
OS Raphanus sativus (Raphan).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Raphanus.
KW NCBI_TaxID=3726;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, AVIGNON; TISSUP=Seed;
RC RAYNAL M., DELIGNY D., COOKE R., DELSENY M.;
RL "Characterization of a radish nuclear gene expressed during late seed
RL maturation."
FT Plant Physiol. 91:829-835(1989).
CC -1- FUNCTION: THIS PROTEIN MAY PLAY A ROLE IN EQUIPPING THE SEED FOR
CC SURVIVAL, MAINTAINING A MINIMAL LEVEL OF HYDRATION IN THE DRY
CC ORGANISM AND PREVENTING THE DENATURATION OF CYTOPLASMIC
CC COMPONENTS, OR MAY PLAY A ROLE DURING INHIBITION BY CONTROLLING
CC WATER UPTAKE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
CC -1- DEVELOPMENTAL STAGE: EXPRESSED DURING LATE SEED MATURATION.
CC -1- INDUCTION: BY ASCSIC ACID (ABA) (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X15440; CAA33479.1; .
DR EMBL; M31978; AAA33668.1; .
DR PIR; S04884; S04884.
DR InterPro; IPR000389; .
DR Pfam; PF00477; seed.protein.1.
DR PROSITE; P500431; SMALL_HYDR_PLANT_SEED; 1.
KW Seed; Seed embryo.
SQ SEQUENCE 83 AA; 8981 MW; A930FB3FC855172 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 VFGG 4
|||||
21 VFGG 24

RESULT 8
ID PTSN_PSEPU STANDARD; PRT; 89 AA.
AC P33670;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE NITROGEN REGULATORY IIA PROTEIN (EC 2.7.1.69) (ENZYME IIA-NTR)
DE (PHOSPHOTRANSFERASE ENZYME II, A COMPONENT) (FRAGMENT).
GN PTSN.
OS Pseudomonas putida.
CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
CC Pseudomonas.
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN2100;
RC MEDLINE=90152355; PubMed=2695395;
RC Inouye S., Yamada M., Nakazawa R., Nakazawa T.;
RL "Cloning and sequence analysis of the ntra (rpoN) gene of Pseudomonas

PT Putida "1;
 RL Gene 45145-152(1989).
 CC -1- FUNCTION: SEEMS TO HAVE A ROLE IN LINKING CARBON AND NITROGEN
 CC ASSIMILATION (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: PROTEIN N-PHOSPHOTRANSFERASE + SUGAR =
 CC PROTEIN HISTIDINE + SUGAR PHOSPHATE.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE PTS IIA FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; M24916; ; NOT ANNOTATED_CDS.
 DR HSSP; P31222; 1A6J.
 DR InterPro: IPR002178; -;
 DR Pfam: PF00359; PFS_EIIA_2; 1.
 DR PROSITE; PS00372; PFS_EIIA_2; 1.
 DR Phosphotransferase system; Transferase; Phosphorylation.
 DR MOD_RES; 88 89
 DR NCBI; 88 89
 DR SEQUENCE 89 AA; 9488 MW; 8C8A4C53914E1DD3 CRC64;
 SQ
 Query Match 100.0%; Score 23; DB 1; Length 89;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGG 4
 Db 17 VPGG 20
 RESULT 9
 LSMS HUMAN STANDARD; PRT: 90 AA.
 AC Q9V4Y9;
 DT 01-OCT-2000 (Rel. 40, Created)
 DT 01-OCT-2000 (Rel. 40, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE U6 SNRNP ASSOCIATED SM-LIKE PROTEIN LSM5.
 GN LSM5.
 OS Homo sapiens (human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eumetazoa; Primates; Catarrhini; Hominoidea; Homo.
 CC NCBI_TaxID=9606;
 QY [1]
 SEQUENCE FROM N.A.
 AC TISSUE=Placenta;
 RX MEDLINE=99299196; PubMed=10369684;
 RA Salgado-Garrido J., Bragado-Nilsson E., Kandels-Lewis S., Seraphin B.;
 RT "Sm and Sm-like proteins assemble in two related complexes of deep
 RT evolutionary origin.";
 RL EMO J. 18:3451-3462(1999).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA MEDLINE=9452783; PubMed=10523320.
 RA Muesel I., Brames M., Kasper A., Bach A., Wilm M., Luehmann R.;
 RT "Sm and Sm-like proteins assemble in two related complexes of deep
 RT 3'-end of U6 snRNA, thereby facilitating U4/U6 duplex formation in
 RT vitro.";
 RL EMO J. 18:5789-5802(1999).
 CC -1- FUNCTION: PLAYS A ROLE IN U6 SNRNP ASSEMBLY AND FUNCTION. BINDS TO
 CC THE 3' END OF U6 SNRNA, THEREBY FACILITATING U4/U6 DUPLEX
 CC FORMATION IN VITRO.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SNRNP SN PROTEINS FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AJ238097; CAB45868.1; -;
 DR InterPro: IPR001163; -;
 DR Pfam: PF01423; Sm1.1;
 DR PROSITE; PS00477; seed; protein: 1.
 DR Phosphotransferase system; Transferase; Phosphorylation.
 DR MOD_RES; 0 1
 DR NCBI; 0 1
 DR SEQUENCE 90 AA; 9806 MW; 4B45811B47B054DB CRC64;
 SQ
 Query Match 100.0%; Score 23; DB 1; Length 90;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VPGG 4
 Db 82 VPGG 85
 RESULT 10
 EMB5 MAIZE STANDARD; PRT: 91 AA.
 AC P46517;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE LATE EMBRYOGENESIS ABUNDANT PROTEIN EMB564.
 GN EMB564.
 OS Zea mays (maize).
 CC Eukaryota; Eumetazoa; Embryophyta; Tracheophyta; Spermatophyta;
 CC Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae;
 CC Andropogoneae; Zea.
 CC NCBI_TaxID=4577;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. CO255; TISSUE=Embryo;
 RX MEDLINE=91316231; PubMed=1830496;
 RA Williams B., Tsang A.;
 RT "A maize gene expressed during embryogenesis is abscisic
 RT acid-inducible and highly conserved.";
 RL Plant Mol Biol Rep 13(1):1-5(1996).
 CC -1- FUNCTION: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
 CC FAMILY.
 CC HIGHER PLANT SEED EMERGES. THEY MAY PLAY AN ESSENTIAL ROLE IN
 CC SEED SURVIVAL AND IN CONTROLLING WATER EXCHANGES DURING SEED
 CC DESICCATION AND IMBIBITION.
 CC -1- INDUCTION: BY ABSICISIC ACID AND OSMOTIC STRESS.
 CC -1- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
 CC FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL; X55388; CAA35063.1; -;
 DR PIR; S16249; S16249.
 DR PIR; S16249; S16249.
 DR InterPro: IPR000389; -;
 DR Pfam: PF00477; seed; protein: 1.
 DR PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
 DR Seed embryo; Seed; Multigene family.
 KW SEQUENCE 91 AA; 9684 MW; C78517500A7F675C CRC64;
 SQ

Query Match 100.0%; Score 23; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 22 VPGG 25

RESULT 11
 EMBL_PICOL STANDARD; PRT; 91 AA.
 ID EMBL_PICOL
 AC Q40864;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE EM-LIKE PROTEIN
 OS Picea glauca (White spruce)
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Eukaryota; Viridiplantae; Coniferales; Pinaceae; Picea.
 OX NCBI_TAXID=3330;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dong J.Z., Dunstan D.I.;
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
 CC !- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on
 CC the use of this entry, except that the source must be acknowledged. This
 CC entry has been modified to remove long stretches of sequence that
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L42464; AAC5366.1;
 CC InterPro; IPR000389;
 CC Pfam; PF00477; seed_protein; 1.
 CC PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
 CC SEQUENCE 91 AA; 9771 MW; 8FD9D4F41ED0A195E CRC64;

Query Match 100.0%; Score 23; DB 1; Length 91;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 21 VPGG 24

RESULT 12
 EMBL_PICOL STANDARD; PRT; 92 AA.
 ID EMBL_PICOL
 AC Q02973;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE EM-LIKE PROTEIN GEA6.
 OS EM6 OR ATEM6 OR D19H OR ATZG40170.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TAXID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLUMBIA;
 RX MEDLINE=93261424; PubMed=8492809;
 RA Gaubier P., Raynal M., Hull G., Huestis G.M., Grellet F.,

Arenas C., Pages M., Delseny M.;
 "Two different Em-like genes are expressed in Arabidopsis thaliana
 seeds during maturation";
 Mol. Gen. Genet. 238:409-418(1993).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, LANOSBERG ERECTA;
 RX MEDLINE=93261425; PubMed=8492808;
 RA "The seed protein GEA6 is a member of the Em-like protein family.
 P"Abasic acid-insensitive mutations provide evidence for stage-
 specific signal pathways regulating expression of an Arabidopsis late
 embryogenesis-abundant (lea) gene";
 Mol. Gen. Genet. 238:401-408(1993).
 [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99225673; PubMed=10207155;
 RA Wang M.L., Belmonte S., Kim U., Dolan M., Morris J.W., Goodman H.M.;
 "A cluster of ABA-regulated genes on Arabidopsis thaliana BAC
 T07M07";
 Genome Res. 9:325-333(1999).
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV, COLDWATER;
 RX MEDLINE=20083487; PubMed=10517197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Watson T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Unayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana";
 Nature 402:761-768(1999).
 CC !- FUNCTION: IT IS THOUGHT TO PROVIDE PROTECTION FOR THE CYTOPLASM
 CC DURING THE DESICCATION STAGE OF EMBRYO DEVELOPMENT.
 CC !- TISSUE SPECIFICITY: PRESENT ONLY IN NEARLY DRY AND DRY SEEDS.
 CC !- INDUCTION: BY ABSCISIC ACID (ABA).
 CC !- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on
 CC the use of this entry, except that the source must be acknowledged. This
 CC entry has been modified to remove long stretches of sequence that
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; 211923; CAA77981.1;
 CC EMBL; 211924; CAA77982.1;
 CC EMBL; X66023; CAA46821.1;
 CC EMBL; AF085279; AAC25932.1;
 CC EMBL; AC018721; AAF18731.1;
 CC InterPro; IPR000389;
 CC Pfam; PF00477; seed_protein; 1.
 CC PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
 CC SEEDSITE; MULTISPEC FAMILY
 CC SEQUENCE 92 AA; 9934 MW; D35A7879AF73706 CRC64;

Query Match 100.0%; Score 23; DB 1; Length 92;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 21 VPGG 24

RESULT 13
 EMBL_DAUCA
 ID EMBL_DAUCA STANDARD; PRT; 92 AA.


```

AC P17639;
DT 01-AUG-1990 (Rel. 15, Created)
DI 01-AUG-1990 (Rel. 15, Last sequence update)
DI 01-OCT-1996 (Rel. 34, Last annotation update)
DE EMB-1 PROTEIN.
GN EMB-1.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC euasterids II; Apiales; Apiaceae; Daucus.
NCBI_TaxID=4039;
[1] SEQUENCE FROM N.A.
RX STRAIN=CV; DANVERS; TISSUE=Embryo;
RX MEDLINE=90251475; PubMed=2339072;
RX Wurtel T.H., Wurtel E.S., Nikolai B.J.;
RX "Sequence of EMB-1, an mRNA accumulating specifically in embryos of
RX carrot.",
RX Nucleic Acids Res. 18:2826-2826(1990).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=CV; DANVERS; TISSUE=Embryo;
RP MEDLINE=94151426; PubMed=8108498;
RP Wurtel E.S., Wang H., Durgarian S., Nikolai B.J., Ulrich T.H.;
RP "Characterization of a gene that is expressed early in somatic
RP differentiation of a carrot root.",
RP Plant Physiol. 102:303-312(1993).
CC -1- DEVELOPMENTAL STAGE: EMBRYOS.
CC -1- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to: license@isb-sib.ch).
DR EMBL; X60131; CAA42717.1;
DR EMBL; X17608; CAA35610.1;
DR PIR; S10010; S10010.
DR PIR; J02273; J02273.
DR InterPro; IPR000389;
DR Pfam; PF00477; seed protein; 1.
DR PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
KW Seed; Seed embryo.
SQ SEQUENCE 92 AA; 9917 MW; 13809FA50F5FA7FA CRC64;

Query Match 100.0%; Score 23; DA 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VFEG 4
DB 21 VFEG 24
|||||
RESULT 14
ID LEIO_HELAN STANDARD; PRT; 92 AA.
AC P32411; 1993 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 10 KDA LATE EMBRYOGENESIS ABUNDANT PROTEIN (DS10).
OS Helianthus annuus (Common sunflower).
OC Eukaryota; Viridiplantae; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Asteridae;
OC euasterids II; Asterales; Asteraceae; Asteroidae; Heliantheae;
OC Helianthus.
NCBI_TaxID=4232;

```

```

RN SEQUENCE FROM N.A.
RP STRAIN=CV; SUNWED; TISSUE=Dry seed;
RP MEDLINE=92353387; PubMed=1385536;
RA Almoquera C., Jordano J.;
RT "Developmental and environmental concurrent expression of sunflower
RT dry-seed-stored low-molecular-weight heat-shock protein and Lea
RT mRNAs.",
RL Plant Mol. Biol. 19:781-792(1992).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=CV; SUNWED; TISSUE=Cotyledon;
RP Submitted (JAN-1998) to the EMBL/GenBank/DBSJ databases.
[3]
RN SEQUENCE FROM N.A.
RX MEDLINE=99190608; PubMed=10092187;
RX Prieto-Dapena P., Almoquera C., Rojas A., Jordano J.;
RX "Seed-specific expression patterns and regulation by AB13 of an
RX unusual late embryogenesis abundant gene in sunflower.",
RL Plant Mol. Biol. 39:615-627(1999).
CC -1- FUNCTION: LEA PROTEINS ARE LATE EMBRYONIC PROTEINS ABUNDANT IN
CC HIGHER PLANT SEED EMBRYOS. THEY MAY PLAY AN ESSENTIAL ROLE IN
CC PROTECTING THE EMBRYO DURING WATER EXCHANGES DURING SEED
CC DEHYDRATION AND IMBIBITION.
CC -1- TISSUE SPECIFICITY: MAXIMALLY EXPRESSED IN DRY SEEDS. ALSO
CC PRESENT IN MID-MATURATION EMBRYOS.
CC -1- INDUCTION: BY ABSCISIC ACID, OSMOTIC STRESS AND HEAT SHOCK.
CC -1- SIMILARITY: BELONGS TO THE SMALL HYDROPHILIC PLANT SEED PROTEIN
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to: license@isb-sib.ch).
DR EMBL; X59699; CAA42220.1;
DR EMBL; AJ224116; CAA11834.1;
DR InterPro; IPR000389;
DR Pfam; PF00477; seed protein; 1.
DR PROSITE; PS00431; SMALL_HYDR_PLANT_SEED; 1.
KW Seed embryo; Seed.
SQ SEQUENCE 92 AA; 10037 MW; ED612FAACE72BAFF CRC64;

Query Match 100.0%; Score 23; DB 1; Length 92;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VFEG 4
DB 31 VFEG 34
|||||
RESULT 15
ID RL44_PALMA STANDARD; PRT; 92 AA.
AC P32411; 1993 (Rel. 37, Created)
DT 01-OCT-1993 (Rel. 37, Last sequence update)
DT 01-OCT-1993 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S RIBOSOMAL PROTEIN L44S (LA) (HLA).
GN RPL44E.
OS Halocaula marismortui (Halobacterium marismortui).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halocaula.
NCBI_TaxID=2238;
[1]
RN SEQUENCE.
RP MEDLINE=93277953; PubMed=8504167;
RA Bergmann U., Wittmann-Liebold B.;

```


No.	Score	Match	Length	DB	ID	Description
1	23	100.0	20	13	Q9PRW3	Q9PRW3 crotalus at
2	23	100.0	25	2	Q46499	Q46499 desulfuovib
3	23	100.0	27	14	Q89142	Q89142 hepatitis c
4	23	100.0	28	10	Q9S886	Q9S886 citrui sine
5	23	100.0	30	6	Q9NZH3	Q9NZH3 bos taurus
6	23	100.0	31	4	Q9SMHO	Q9SMHO homo sapien
7	23	100.0	41	2	Q93920	Q93920 streptomyces
8	23	100.0	42	8	Q9ZT08	Q9ZT08 lactococcus
9	23	100.0	44	6	Q9ZT08	Q9ZT08 lactococcus
10	23	100.0	44	8	Q9ZF23	Q9ZF23 arycollagus
11	23	100.0	45	6	Q9T0R1	Q9T0R1 bos taurus
12	23	100.0	51	10	Q9SNE3	Q9SNE3 gossypium s
13	23	100.0	56	2	Q9XVW8	Q9XVW8 streptomyces
14	23	100.0	57	7	Q29978	Q29978 homo sapien
15	23	100.0	61	2	Q53465	Q53465 streptomyces
16	23	100.0	63	10	Q49576	Q49576 arabidopsis
17	23	100.0	67	2	Q918X0	Q918X0 streptomyces
18	23	100.0	70	2	Q9PK17	Q9PK17 streptomyces
19	23	100.0	70	2	Q52113	Q52113 acinetobact

DT 01-NOV-1996 (REMBREL. 01, Last annotation update)
DE PERIPLASMIC (FRAGMENT).
OS Desulfovibrio desulfuricans.

US DESTROYED RESULTS.

100

OC Bacteria; Proteobacteria; delta subdivision; Desulfovibrio.

OK NCBI_TaxID=876;
RN [1] SEQUENCE FROM N.A.
PC SPRAIN-46;
RA Lin S.K., Lee S.J., Kim B.H.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U49192; AAA91808.1; -.
FT NON_TER 1 1
FT NON_TER 20 20
SQ SEQUENCE 20 AA; 1896 MW; E93207D26C22998 CRC64;

Query Match 100.0%; Score 23; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 VFGG 4
DB 3 VFGG 6

RESULT 3
ID Q89142 PRELIMINARY: PRT; 27 AA.
AC Q89142;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE BASIC CHITINASE (EC 3.2.1.14) (FRAGMENT).
OS Hepatitis C virus
OC Hepatitis C virus
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-18-2;
RX MEDLINE=98378035; PubMed=9714233;
RA Kato N., Ikeda M., Sugiyama K., Mizutani T., Tanaka T., Shimotohno K.;
RT Hepatitis C virus population dynamics in human lymphocytes and
RT hepatocytes in transgenic mice.
RL GenBank: U49192; AAA91808.1; -.
DR EMBL: AB007716; BA32637.1; -.
DR INTERPRO: IPR002531; -.
DE PFAM: PF01560; HCV_NSI: 1.
FT NON_TER 1 1
FT NON_TER 27 27
SQ SEQUENCE 27 AA; 2861 MW; BE1F4D9A2AB0A25 CRC64;

Query Match 100.0%; Score 23; DB 14; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e+02; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 VFGG 4
DB 4 VFGG 7

RESULT 4
ID Q98886 PRELIMINARY: PRT; 28 AA.
AC Q98886;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BASIC CHITINASE (EC 3.2.1.14) (FRAGMENT).
OS Citrus sinensis (Sweet orange).
OC Rutaceae; Rutaceae; Citrus.
OC Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=2711;
RN [1]

RP SEQUENCE.
RX MEDLINE=97085076; PubMed=8931350;
RA Kaya R., McDonald R.F., Hearn C.J., McDonald R.F.,
RA Kaya R., McDonald R.F.;
PC "Characterization of seven basic endochitinases isolated from cell
RT cultures of Citrus sinensis (L.)";
RL Planta 200:289-295(1996).
DR HSSE: P25871; LAUN.
DR INTERPRO: IPR001938; -.
DR PFAM: PF00314; thaumatocin; 1.
DR PRINTS: PR00347; THAUMATIN.
SQ SEQUENCE 28 AA; 3041 MW; FE27A9A1AD305AF5 CRC64;

Query Match 100.0%; Score 23; DB 10; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.9e+02; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 VFGG 4
DB 18 VFGG 21

RESULT 5
ID Q9NZH3 PRELIMINARY: PRT; 30 AA.
AC Q9NZH3; 2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE BUCENTAUROS (FRAGMENT).
GN BCNT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Buthera; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Ruminantia; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Takahashi I., Seki M., Takahashi I., Takeda H., Sugimoto Y.;
RT "Genomic organization of bovine bcnt that contains LINE-derived region in
RT its protein";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Takahashi I., Nobukuni T., Ohmori H., Kobayashi M., Tanaka S.,
RA Ohshima K., Okada N., Kasui T., Hashimoto K., Iwashita S.;
RT "Existence of a bovine LINE repetitive insert that appears in the cDNA
RT of bovine protein BCNT in ruminant, but not in human, genomes.";
RL Gene 211:387-394(1998).
DR EMBL: AB033999; BA93548.1; -.
FT NON_TER 1 1
FT NON_TER 30 30
SQ SEQUENCE 30 AA; 2951 MW; D32134B9AF2A332F CRC64;

Query Match 100.0%; Score 23; DB 6; Length 30;
Best Local Similarity 100.0%; Pred. No. 2.1e+02; Mismatches 0; Gaps 0;
Matches 4; Conservative 0; Indels 0;
QY 1 VFGG 4
DB 11 VFGG 14

RESULT 6
ID Q9DHHO PRELIMINARY: PRT; 31 AA.
AC Q9DHHO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE TAU PROTEIN (FRAGMENT).

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE-AGED BRAIN;
RX MEDLINE=89193714; PubMed=2495000;
Mori H., Hamada Y., Kawaguchi M., Honda T., Kondo J., Ihara Y.;
RT "A distinct form of tau is selectively incorporated into Alzheimer's
RT paired helical filaments";
RL Biochem. Biophys. Res. Commun. 159:1221-1226(1989).
DR INTERPRO: IP001084; -;
DR EMBL: M25298; AAA57264.1; -;
DR FRAM: PR00418; tubulin-binding; 1.
DR PROSITE: PS00229; TAU_MAF; 1.
DR NON_TER 31 31
FT NON_TER 31 31
SQ SEQUENCE 31 AA; 3365 MW; 5C52909A643AC9E CRC64;

Query Match 100.0%; Score 23; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 26 VPGG 29

RESULT 7
ID Q9X9Z0 PRELIMINARY; PRT; 41 AA.
AC Q9X9Z0;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE DE PUTATIVE TWO COMPONENT SENSOR KINASE (FRAGMENT).
GN SC15.10C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteriae; Actinobacteridae;
OC Streptomyces; Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]

SEQUENCE FROM N.A.
STRAIN=A3(2);
Seeger K., Harris D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb streptomycetes coelicolor A3(2) chromosome";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]

SEQUENCE FROM N.A.
ID Q9X9Z0 PRELIMINARY; PRT; 41 AA.
AC Q9X9Z0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CHAPERONIN-T-COMPLEX POLYPEPTIDE 1 HOMOLOG (FRAGMENTS).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
Redenbach M., Kleser H.M., Denaparte D., Elchner A., Cullum J.,
Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb streptomycetes coelicolor A3(2) chromosome";
RL EMBL: AL070332; G334529.1; -;
DR INTERPRO: IP000410; -;
DR PRINTS: PR00344; BCTRLSENSOR.
KW Kinase.
FT NON_TER 1 1
SQ SEQUENCE 41 AA; 4221 MW; ABCC81778A9A0D CRC64;

Query Match 100.0%; Score 23; DB 2; Length 41;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 38 VPGG 41

RESULT 8
ID Q92FC8 PRELIMINARY; PRT; 44 AA.
AC Q92FC8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE THIOREDOXIN REDUCTASE (FRAGMENT).
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ILL403;
RA Aleksandrak T., Kowalczyk M., Kok J., Bardowski J.;
RT "Dual role of CcpA protein in regulation of sugar metabolism in
RT Lactococcus lactis";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF06673; AAC96331.1; -;
DR INTERPRO: IP000105; -;
DR PRINTS: PR000750; -;
DR PRINTS: PR00419; ADXCTPASE.
DR PRINTS: PR00469; PHOTOTASEII.
DR PRINTS: PR01001; FADG3PDH.
FT NON_TER 44 44
SQ SEQUENCE 44 AA; 4503 MW; 5B9AF8B3C141140C CRC64;

Query Match 100.0%; Score 23; DB 2; Length 44;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
DB 37 VPGG 40

RESULT 9
ID Q9TSD3 PRELIMINARY; PRT; 44 AA.
AC Q9TSD3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE CHAPERONIN-T-COMPLEX POLYPEPTIDE 1 HOMOLOG (FRAGMENTS).
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94089752; PubMed=7903455;
Vanderkerke H., Yen Hoogs G., Gao Y., Melki R., Cowan N.J.,
PA "Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and
RT seven related subunits";
RL Proc. Natl. Acad. Sci. U.S.A. 90:11975-11979(1993).
FT NON_TER 1 1
FT NON_TER 18 19
FT NON_TER 44 44
SQ SEQUENCE 44 AA; 4458 MW; E631D615DD0DF2 CRC64;

Query Match 100.0%; Score 23; DB 5; Length 44;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 27 VPGG 30

RESULT 10
 Q97213
 ID Q97213 PRELIMINARY; PRT; 44 AA.

AC Q97213; 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 13, Last annotation update)
 DE CHAPERONIN-60 157 FRAGMENT (FRAGMENTS).
 OS Brassica napus (Rape).
 OG Mitochondrion.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Brassica.
 NCBI_TaxID=3708;
 [1]

RP SEQUENCE.
 RX MEDLINE=94302168; PubMed=7913238;
 RA Cloney L.P., Bekasov D.R., Feist G.L., Lane W.S., Hemmingsen S.N.;
 RT "Strasburg, Pa. and the mitochondrial chaperonin-60 proteins
 RL Plant Physiol. 105:233-241 (1994).
 DR RSP; P06139; 1SER.
 DE INTERPRO: IPR001844; -.
 DE INTERPRO: IPR002423; -.
 DR PFAM: PF00118; cpm60_TCP1; 1.
 DR PROSITE: PS00296; CHAPERONIN_CPM60; 1.
 FT NON_TER 1 1
 FT NON_CONS 23 24
 FT NON_TER 44 44
 SQ SEQUENCE 44 AA; 4341 MW; 8443D0C9452AF4F3 CRC64;

Query Match 100.0%; Score 23; DB 8; Length 44;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 32 VPGG 35

RESULT 11
 Q971

ID Q971 PRELIMINARY; PRT; 45 AA.

AC Q971; 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TRENBLrel. 14, Last annotation update)
 DE CHAPERONIN (FRAGMENTS).
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 [1]
 RP SEQUENCE.
 RX MEDLINE=94089752; PubMed=7903455;
 RA Rommelaere H., Van Troys M., Gao Y., Melki R., Cowan N.J.,
 RA Vandekerckhove J., Ape C.;
 RT "Eukaryotic cytosolic chaperonin contains t-complex polypeptide 1 and
 seven related subunits";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:11975-11979 (1993).
 RN [2]

RP SEQUENCE.
 RX MEDLINE=9309850; PubMed=1361170;
 RA Frydman J., Wilmesgren E., Erdjument-Bromage H., Hall J.S., Tempst P.,
 RA Hartl P.O.;
 RT "Function in protein folding of Tric, a cytosolic ring complex
 RI consisting of seven related subunits";
 RI EMBO J. 11:4767-4778 (1992).
 DR INTERPRO: IPR002423; -.
 DR PFAM: PF00118; cpm60_TCP1; 1.
 FT NON_TER 1 1
 FT NON_CONS 20 21
 FT NON_TER 45 45
 SQ SEQUENCE 45 AA; 4962 MW; 1573AB4E23BCD9 CRC64;

Query Match 100.0%; Score 23; DB 6; Length 45;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 29 VPGG 32

RESULT 12
 Q95F3

ID Q95F3 PRELIMINARY; PRT; 51 AA.
 AC Q95F3; 2000 (TRENBLrel. 13, Created)
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TRENBLrel. 13, Last annotation update)
 DE MICROSOMAL OMEGA5 DESATURASE ENZYME (FRAGMENT).
 GN PAD2.
 OS Gossypium stocksii.
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Malvales; Malvaceae; Gossypium.
 NCBI_TaxID=47602;
 [1]
 RP SEQUENCE FROM N.A.
 RA Prohasky C.L., Liu Q., Green A.G., Marshall D.R., Singh S.P.;
 RT "Microsomal omega5 desaturase: iontophoresis contribute to our
 RT understanding of reticulate evolution in Gossypium (Malvaceae) and the
 RL evolution of the Australian Gossypium species";
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ244917; CAB59280.1; -.
 FT NON_TER 1 1
 FT CHAIN 1 >51
 FT NON_TER 51 51
 SQ SEQUENCE 51 AA; 5728 MW; B3103EFF7594B8A3 CRC64;

Query Match 100.0%; Score 23; DB 10; Length 51;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 2 VPGG 5

RESULT 13
 Q9XV8

ID Q9XV8 PRELIMINARY; PRT; 56 AA.
 AC Q9XV8; 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE HYPOTHETICAL 6.4 KDA PROTEIN.
 GN SC17.32.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.

OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RC Sequences involved in growth-phase-dependent expression and glucose
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RC Sequences involved in growth-phase-dependent expression and glucose
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 DR EMBL; AL096743; CAB46415.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 56 AA; 6392 MW; CBE7F4D6BFU413B8 CRC64;

Query Match 100.0%; Score 23; DB 2; Length 56;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 5 VPGG 8

RESULT 14
 Q29978 PRELIMINARY; PRT; 57 AA.
 AC Q29978;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 07, Last annotation update)
 DE LYMPHOCTE ANTIGEN (FRAGMENT).
 HLA-DQB1.
 Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Meyer C.G.;
 RL Tissue Antigens 0:0-0(0).
 DR EMBL; M86226; AAH5692.1; -;
 KW HMG.
 FT NON_TER 1 1
 SQ SEQUENCE 57 AA; 5594 MW; DD964A059393C73 CRC64;

Query Match 100.0%; Score 23; DB 7; Length 57;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 27 VPGG 30

RESULT 15
 Q53455 PRELIMINARY; PRT; 61 AA.
 AC Q53455;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE CRE 5; OF AAL (FRAGMENT).
 OC Streptomyces (Actinobacteria); Actinobacteriales; Actinobacteriaceae; Streptomyces.
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94297727; PubMed-8025672;
 RA Virolle M.J., Gagnat J.;
 RT "Sequences involved in growth-phase-dependent expression and glucose
 RL repression of a Streptomyces alpha-amylase gene.";
 RL Microbiology 140:1059-1067(1994).
 DR EMBL; S73706; AB31287.1; -;
 FT NON_TER 1 1
 SQ SEQUENCE 61 AA; 6569 MW; 655F08804E525322 CRC64;

Query Match 100.0%; Score 23; DB 2; Length 61;
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VPGG 4
 DB 58 VPGG 61

Search completed: April 24, 2001, 16:40:39
 Job time: 429 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:38:20 ; Search time 115.25 Seconds
(without alignments)
14.880 Million cell updates/sec

Title: US-09-340-736-8
Perfect score: 174
Sequence: 1 GSGVGLGYGGLGYGGLGYGGLY 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Archived: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04

Maximum Match 1004

Listing first 45 summaries

Database : A-Genesec 0401.*

- 1: /SIDSL/gcgdata/genesecp/AA1980.DAT.*
- 2: /SIDSL/gcgdata/genesecp/AA1981.DAT.*
- 3: /SIDSL/gcgdata/genesecp/AA1982.DAT.*
- 4: /SIDSL/gcgdata/genesecp/AA1983.DAT.*
- 5: /SIDSL/gcgdata/genesecp/AA1984.DAT.*
- 6: /SIDSL/gcgdata/genesecp/AA1985.DAT.*
- 7: /SIDSL/gcgdata/genesecp/AA1986.DAT.*
- 8: /SIDSL/gcgdata/genesecp/AA1987.DAT.*
- 9: /SIDSL/gcgdata/genesecp/AA1988.DAT.*
- 10: /SIDSL/gcgdata/genesecp/AA1989.DAT.*
- 11: /SIDSL/gcgdata/genesecp/AA1990.DAT.*
- 12: /SIDSL/gcgdata/genesecp/AA1991.DAT.*
- 13: /SIDSL/gcgdata/genesecp/AA1992.DAT.*
- 14: /SIDSL/gcgdata/genesecp/AA1993.DAT.*
- 15: /SIDSL/gcgdata/genesecp/AA1994.DAT.*
- 16: /SIDSL/gcgdata/genesecp/AA1995.DAT.*
- 17: /SIDSL/gcgdata/genesecp/AA1996.DAT.*
- 18: /SIDSL/gcgdata/genesecp/AA1997.DAT.*
- 19: /SIDSL/gcgdata/genesecp/AA1998.DAT.*
- 20: /SIDSL/gcgdata/genesecp/AA1999.DAT.*
- 21: /SIDSL/gcgdata/genesecp/AA2000.DAT.*
- 22: /SIDSL/gcgdata/genesecp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	DB ID	Description
1	109	62.6	123	17	R95138	Silk like protein
2	109	62.6	123	22	B64033	SELP2-SLFP functio
3	109	62.6	139	17	R95139	Silk like protein
4	109	62.6	139	22	B64038	SELP3-SLFP functio
5	96	55.2	135	21	B19198	Amino acid sequenc
6	96	55.2	135	21	B19196	Amino acid sequenc
7	95.5	54.9	294	9	P82484	Tropoelastin, Gal
8	91	52.3	126	21	B19197	Amino acid sequenc
9	91	52.3	521	18	M36054	Mouse occludin pro
10	89	51.1	646	18	M27178	Nephila clavipes s
11	88.5	50.9	93	21	G24552	Arabiopsis thalia

12	88.5	50.9	104	21	G24551	Arabiopsis thalia
13	88.5	50.9	133	21	G30160	Arabiopsis thalia
14	88.5	50.9	133	21	G52833	Arabiopsis thalia
15	88.5	50.9	158	21	G52832	Arabiopsis thalia
16	88.5	50.9	158	21	G52832	Arabiopsis thalia
17	88.5	50.9	181	21	G52832	Arabiopsis thalia
18	88.5	50.9	253	21	G36620	Arabiopsis thalia
19	88.5	50.9	253	21	G36620	Arabiopsis thalia
20	88.5	50.9	273	21	G36619	Arabiopsis thalia
21	88.5	50.9	309	21	G36618	Arabiopsis thalia
22	88	50.6	604	16	R99057	Spider dragline va
23	88	50.6	606	16	R99055	Spider dragline va
24	88	50.6	606	20	Y40101	Polymer of an anal
25	88	50.6	606	20	Y40102	Polymer of an anal
26	88	50.6	651	20	Y40057	Spider silk protei
27	88	50.6	718	12	KL3308	Nephila clavipes s
28	88	50.6	718	19	W53346	Nephila clavipes s
29	88	50.6	718	21	W53346	Nephila clavipes s
30	86.5	49.7	339	21	G17435	Arabiopsis thalia
31	86.5	49.7	339	21	G17435	Arabiopsis thalia
32	86.5	49.7	339	21	G20657	Arabiopsis thalia
33	86.5	49.7	421	21	G17434	Arabiopsis thalia
34	86.5	49.7	421	21	G20656	Arabiopsis thalia
35	86.5	49.7	612	21	G17433	Arabiopsis thalia
36	86.5	49.7	612	21	G20655	Arabiopsis thalia
37	85	48.9	176	21	G11337	Arabiopsis thalia
38	85	48.9	187	21	G11336	Arabiopsis thalia
39	85	48.9	206	21	G11335	Arabiopsis thalia
40	85	48.9	232	16	W05704	Glycine-rich repea
41	85	48.9	232	16	W05704	Glycine-rich repea
42	85	48.9	232	19	W73137	EGF-like factor
43	85	48.9	611	20	Y28843	Esten Barr Virus
44	85	48.9	641	21	Y98836	Esten Barr Virus
45	84.5	48.6	64	21	G16184	Arabiopsis thalia

ALIGNMENTS

RESULT 1
ID R95138 standard; Protein; 123 AA.
XX R95138;
XX R95138;
DT 03-FEB-1997 (first entry)
XX
DE SILK like protein (SELP)2-SLFP.
XX
KW Polymer; repeat unit; natural fibroin; intervening oligopeptide;
KW fibre; film; membrane; emulsion; coating; silk like protein;
KW specific binding material; catalyst; purification agent; composite;
KW laminate; adhesive; cell growth surface; affinity column;
KW biological material support; wound dressing; in vivo prothesis.
XX Synthetic.
XX US5514581-A.
XX
PD 07-MAY-1996.
XX
PF 04-NOV-1986; 86US-0927258.
XX
PR 06-NOV-1990; 90US-0609716.
PR 04-NOV-1985; 86US-0927258.
PR 23-OCY-1987; 87US-014648.
PR 07-NOV-1988; 86US-026543.
PR 07-NOV-1989; 89US-0359315.
XX
PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
XX
PI Cappello J, Ferrari FA;
XX

DR WPI; 1996-238772/24.
XX DNA encoding protein contg. repeated fibroin derived segments -
PT linked oligopeptide with cell adhesion properties useful, e.g. in
PT wound dressings
XX
XX Example 3; Columns 127-128; 71pp; English.
XX Novel DNA sequence encodes a polymer comprising segments of
CC repeating units of 3-9 amino acids from natural fibroin, i.e. the
CC present silk like protein (SELP)2-SLPF sequence, able to assemble
CC into aligned structures formable into articles. The polymer
CC comprises at least 2 segments joined by an unaligned intervening
CC oligopeptide, other than the repeating unit.
CC The polymer can be used to make fibres, films, membranes,
CC materials, catalysts, purific agents, composites, laminates,
CC adhesives, cell growth surfaces, affinity columns and supports for
CC biological materials. Typical applications include wound
CC dressings, and in vivo protheses. The polymer produces articles
CC with good mechanical properties, and the intervening oligopeptide
CC can provide a ligand for binding a mol., antibody, etc., or a
CC chemically reactive site for coupling to proteins, etc..
SQ Sequence 123 AA;
Query Match 62.6%; Score 109; DB 17; Length 123;
Best Local Similarity 62.1%; Pred. No. 4.5e-06;
Matches 18; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
OY 1 GGLGVLGVLGVLGVLGVLGVLGVLGVLG 29
DB 35 GSVGVGVGVGVGVGVGVGVGVGVGVGVGV 63
RESULT 2
ID B64033 standard; Protein; 123 AA.
AC B64033;
DT 19-MAR-2001 (first entry)
DE SELP2-SLPF functional polymer amino acid sequence SEQ ID 103.
XX Proteinaceous polymer; repeat unit; structural polymer; coating; film;
KW fibre; membrane; adhesive; emulsion; laminate; keratin; collagen.
XX Synthetic.
OS US6140072-A.
PD 31-OCT-2000.
PF 07-JUN-1995; 95US-0475411.
PR 06-NOV-1990; 90US-0609716.
PR 04-NOV-1986; 86US-0927259.
PR 29-OCT-1987; 87US-0114518.
PR 09-NOV-1988; 88US-0269429.
XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
PA Cappello J, Ferrari FA;
PI WPI; 2001-048958/06.
DR
XX New DNA encoding a polymer with strands of repeating units of natural
PT protein joined by intervening oligopeptide for producing high molecular
PT weight polymers of amino acids
XX
XX Example 2; Column 49-50; 73pp; English.

XX This invention relates to DNA encoding a proteinaceous polymer. The
CC polypeptide strands of repeating units of a natural protein, capable
CC of assembling into aligned structures with at least 2 strands joined by
CC an intervening oligopeptide other than the repeating units. The
CC intervening oligopeptide is unaligned and the polymer has individual
CC strands of the same or different repeating units. The DNA is useful for
CC producing high molecular weight polymers of amino acids based on
CC biologically and chemically active structural polymers. These polymers
CC may be used to provide a variety of structures for different purposes,
CC and to produce articles including coatings, or other (non)structural
CC components, e.g. fibres, films, membranes, adhesives or emulsions, or
CC with other compounds and/or compositions to form composites or laminates.
CC Reptide sequences of 3-9 amino acids from natural fibroin can be
CC used to produce aligned structures. Oligonucleotide sequences
CC of 23370 - 23386 and amino acid sequences B63992 - B64002 are used in the
CC construction of SUP and PCB-SLP polymers. Oligonucleotide sequences
CC of 23387 - 23397 and amino acid sequences B64003 - B64008 are used in the
CC construction of CLP (collagen like protein) polymers. Oligonucleotide
CC sequences 23398 - 23409 and amino acid sequences B64009 - B64014 are
CC used in the construction of keratin polymers. Proteins and peptides
CC represented by sequences B64015 - B64049 are examples of polymers of the
XX invention.
XX Sequence 123 AA;
Query Match 62.6%; Score 109; DB 22; Length 123;
Best Local Similarity 62.1%; Pred. No. 4.5e-06;
Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
OY 1 GGLGVLGVLGVLGVLGVLGVLGVLGVLG 29
DB 35 GSVGVGVGVGVGVGVGVGVGVGVGVGVGV 63
RESULT 3
ID R95139 standard; Protein; 159 AA.
AC R95139;
DT 03-FEB-1997 (first entry)
DE Silk like protein (SELP)3-SLPF.
XX Polymer; repeat unit; natural fibroin; intervening oligopeptide;
KW fibre; film; membrane; emulsion; coating; silk like protein;
KW specific binding material; catalyst; purification agent; composite;
XX laminate; adhesive; cell growth surface; affinity column;
XX biological material support; wound dressing; in vivo prothesis.
OS Synthetic.
PR US5514581-A.
XX
XX 07-MAY-1996.
PD 04-NOV-1986; 86US-0927258.
PF 06-NOV-1990; 90US-0609716.
PR 04-NOV-1986; 86US-0927258.
PR 29-OCT-1987; 87US-0114518.
PR 09-NOV-1988; 88US-0269429.
XX (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.
PA Cappello J, Ferrari FA;
PI WPI; 1996-238772/24.
DR
XX DNA encoding protein contg. repeated fibroin derived segments -

PT linked by oligopeptide with cell adhesion properties useful, e.g. in
 XX wound dressings

XX Example 3; Columns 129-130; 71pp; English.

XX Novel DNA sequence encodes a polymer comprising segments of
 CC repeating units of 3-9 amino acids, natural fibroblast, the
 CC present silk like protein (SLP)-SLP segments, the polymer
 CC into aligned structures formable into articles, the polymer
 CC comprises at least 2 segments joined by an unaligned intervening
 CC oligopeptide, other than the repeating unit.
 CC The polymer can be used to make fibres, films, membranes,
 CC emulsions, coatings, etc., useful as, e.g. specific binding
 CC materials, catalysts, purification agents, composites, laminates,
 CC adhesives, cell growth surfaces, affinity columns and supports for
 CC biological materials. Typical applications include wound
 CC dressings, and in vivo prostheses. The polymer produces articles
 CC with good mechanical properties, and the intervening oligopeptide
 CC can provide a ligand for binding a mol., antibody, etc., or a
 CC chemically reactive site for coupling to proteins, etc..

XX Sequence 159 AA;

Query Match 62.6%; Score 109; DB 17; Length 159;
 Best Local Similarity 62.1%; Pred. No. 5.7e-06;
 Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLYGGGGLGGGLGGGLGGGLGGGLGG 29
 DB 53 GGGGGGGGGGGGGGGGGGGGGGGGG 81

RESULT 4

B64034
 ID B64034 standard; Protein: 159 AA.

XX AC B64034;

XX DT 19-MAR-2001 (first entry)

XX SLF3-SLPF functional polymer amino acid sequence SEQ ID 104.

XX Proteinaceous polymer; repeat unit: structural polymer; coating; film;
 KW fibre; membrane; adhesive; emulsion; laminate; keratin; collagen.

XX Synthetic.

XX US6140072-A.

XX PD 31-OCT-2000.

XX DT 07-JUN-1995; 95US-0475411.

XX PR 06-NOV-1990; 90US-0609716.

XX PR 04-NOV-1986; 86US-0927258.

XX PR 25-OCT-1987; 87US-0114618.

XX PR 09-NOV-1988; 88US-0269429.

XX PA (PROT-) PROTEIN POLYMER TECHNOLOGIES INC.

XX PI Cappelletto J, Ferrari PA;

XX DR WPI, 2001-048559/06.

XX New DNA encoding a polymer with strands of repeating units of natural
 PT protein joined by intervening oligopeptide for producing high molecular
 PT weight polymers of amino acids

XX Example 2; Column 49-50; 73pp; English.

XX This invention relates to DNA encoding a proteinaceous polymer. The
 CC polymer comprises strands of repeating units of a natural protein capable

CC of assembling into aligned structures; with at least 2 strands joined by
 CC an intervening oligopeptide other than the repeating units. The
 CC intervening oligopeptide is unaligned and the polymer has individual
 CC strands of the same or different repeating units. The DNA is useful for
 CC producing high molecular weight polymers of amino acids based on
 CC biologically and chemically active structural polymers. These polymers
 CC may be used to provide a variety of structures for different purposes,
 CC and to produce articles including coatings, other (non)structural
 CC articles, films, fibres, membranes, composites, laminates, or
 CC with other compounds and/or compositions to form composites or laminates.
 CC peptide sequences B63971-B63991 represent monomer sequences which can be
 CC used in the polymers of the invention. Oligonucleotide sequences
 CC F23370 - F23386 and amino acid sequences B63992 - B64002 are used in the
 CC construction of SLP and PCB-SLP polymers. Oligonucleotide sequences
 CC F23387 - F23397 and amino acid sequences B64003 - B64008 are used in the
 CC construction of CUP (collagen like protein) polymers. Oligonucleotide
 CC sequences F23398 - F23409 and amino acid sequences B64009 - B64014 are
 CC used in the construction of keratin polymers. Proteins and peptides
 CC represented by sequences B64015 - B64049 are examples of polymers of the
 CC invention.

XX Sequence 159 AA;

Query Match 62.6%; Score 109; DB 22; Length 159;
 Best Local Similarity 62.1%; Pred. No. 5.7e-06;
 Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLYGGGGLGGGLGGGLGGGLGGGLGG 29
 DB 53 GGGGGGGGGGGGGGGGGGGGGGGGG 81

RESULT 5

B19198
 ID B19198 standard; Protein: 131 AA.

XX AC B19198;

XX DT 19-FEB-2001 (first entry)

XX Amino acid sequence of a bay scallop abductin polypeptide.

XX Bay scallop; abductin; chemomechanical transduction; drug delivery;
 KW tissue adhesive; tissue adhesion; water soluble drug; biomaterial;
 KW fabric; organ prosthesis.

XX Argospecten sp.

XX US6127166-A.

XX PD 03-OCT-2000.

XX PF 03-NOV-1997; 97US-0963168.

XX PR 03-NOV-1997; 97US-0963168.

XX (BAYL/) BAYLEY H.

XX (CAO/) CAO Q.

XX (WANG/) WANG Y.

XX PI Bayley H, Cao Q, Wang Y;

XX DR WPI, 2000-611057/58.

XX N-PSDB; C61379.

XX Abductin nucleic acid molecules useful for expressing abductin

XX proteins and the use of these molecules in the manufacture of drug delivery

XX vehicles for administering water soluble drugs

XX Example 1; Fig 2; 30pp; English.

XX The present sequence represents a bay scallop abductin polypeptide.

XX
XX

106 vsafavag-avag-avag-avagavagavagav 129

PR	04-OCT-1999;	99US-015717.
PR	05-OCT-1999;	99US-015718.
PR	06-OCT-1999;	99US-015725.
PR	07-OCT-1999;	99US-015763.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159283.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159350.
PR	14-OCT-1999;	99US-0159351.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160832.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161962.
PR	28-OCT-1999;	99US-0161963.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity		50.9% Score 88.5; DB
Matches 20; Conservative		Pred. No. 0.0007
		Mismatches 1; Mismatches
QY	1 GCGVGGGAGTG-----GLGVGGGAGTGGGAGTG 29	
Db	55 GGAGGGGAGTGGAGAGAGAGAGAGAGAGGGGGGGGAGGGG 90	
RESULT 12		
G24551	ID	G24551 standard; Protein: 104 AA.
AC	XX	
AC	G24551:	
DE	XX	
DE	17-OCT-2000	(first entry)
DE	Arabidopsis thaliana protein fragment SEQ ID	
KW	Protein identification; signal transduction	
KW	hybridisation assay; genetic mapping; gene ex-	
KW	termination sequence.	
OS	Arabidopsis thaliana.	
XX	XX	
XX	EF1033405-A2.	
PD	06-SEP-2000.	
PF	25-FEB-2000;	2000EP-0301439.
PR	25-FEB-1999;	99US-0121825.
PR	05-MAR-1999;	99US-0122180.
PR	09-MAR-1999;	99US-0122548.
PR	22-MAR-1999;	99US-0122768.
PR	22-MAR-1999;	99US-0122769.
PR	01-APR-1999;	99US-0122785.
PR	01-APR-1999;	99US-0122787.

[illegible]

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:35:25 ; Search time 62.39 Seconds
(without alignments)
9.237 Million cell updates/sec

Title: US-09-340-736-8
Sequence: 1 GGLGVLGGLGGLGGLGGLGGLGGLG 30

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Aligned: 185757 segs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_5/ptodata/2/iaa/5A.COMB.pep:*

2: /cgn2_5/ptodata/2/iaa/5E.COMB.pep:*

3: /cgn2_5/ptodata/2/iaa/6A.COMB.pep:*

4: /cgn2_5/ptodata/2/iaa/6B.COMB.pep:*

5: /cgn2_5/ptodata/2/iaa/PCTUS.COMB.pep:*

6: /cgn2_5/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the best result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query %	ID	Description
1	174	100.0	30	2 US-08-911-364-8	Sequence 8, Appl
2	109	62.6	123	1 US-07-609-716-103	Sequence 103, App
3	109	62.6	123	4 US-08-478-411A-103	Sequence 103, App
4	109	62.6	123	1 US-07-609-716-103	Sequence 103, App
5	109	62.6	159	1 US-07-609-716-103	Sequence 104, App
6	109	62.6	159	4 US-08-478-411A-104	Sequence 104, App
7	109	62.6	159	4 US-08-478-411A-104	Sequence 104, App
8	96	55.2	111	4 US-08-963-168C-15	Sequence 15, Appl
9	96	55.2	116	4 US-08-963-168C-13	Sequence 13, Appl
10	96	55.2	131	4 US-08-963-168C-8	Sequence 8, Appl
11	96	55.2	136	4 US-08-963-168C-6	Sequence 6, Appl
12	93	53.4	745	2 US-09-010-928B-28	Sequence 28, Appl
13	93	53.4	870	2 US-09-010-928B-2	Sequence 2, Appl
14	93	53.4	907	2 US-09-010-928B-4	Sequence 4, Appl
15	93	53.4	907	2 US-09-010-928B-4	Sequence 4, Appl
16	93	53.4	907	2 US-09-010-928B-4	Sequence 4, Appl
17	88	50.6	718	2 US-08-963-168C-7	Sequence 7, Appl
18	88	50.6	718	2 US-08-963-168C-7	Sequence 7, Appl
19	88	50.6	718	2 US-08-963-168C-7	Sequence 7, Appl
20	87.5	50.3	112	4 US-08-963-168C-16	Sequence 16, Appl
21	87.5	50.3	132	4 US-08-963-168C-9	Sequence 9, Appl
22	85	48.9	235	2 US-08-526-190E-1	Sequence 1, Appl
23	82	47.1	341	2 US-08-536-711A-8	Sequence 8, Appl
24	82	47.1	353	2 US-08-536-711A-7	Sequence 7, Appl
25	81	46.6	400	3 US-08-815-190A-17	Sequence 17, Appl
26	79	45.4	1345	2 US-08-977-767-3	Sequence 3, Appl
27	77.5	44.5	832	1 US-08-208-747-2	Sequence 2, Appl

Sequence 2, Appl
Sequence 1, Appl
Sequence 40, Appl
Sequence 41, Appl
Sequence 42, Appl
Sequence 43, Appl
Sequence 44, Appl
Sequence 1, Appl
Sequence 3, Appl
Sequence 18, Appl
Sequence 91, Appl
Sequence 2, Appl
Sequence 27, Appl
Sequence 27, Appl
Sequence 20, Appl
Sequence 3, Appl
Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-911-364-8
; Sequence 8, Application US/08911364
; Patent No. 5969106
; GENERAL INFORMATION
; INVENTOR: ROSENSTEIN, Asst
; APPLICANT: KERN, David W
; TITLE OF INVENTION: SELF-ALIGNING PEPTIDES MODELED ON HUMAN
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-1100
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/911,364
; FILING DATE: 07-AUG-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/023,552
; ATTORNEY/AGENT INFORMATION:
; NAME: Brent, Stephen
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 041082/0104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
US-08-911-364-8

Query Match 100.0%; Score 174; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 5.2e-15;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGLGTYGGLGGLGGLGGLGGLGGLGGLG 30
 |||||
 DB 1 GGLGTYGGLGGLGGLGGLGGLGGLGGLG 30

RESULT 2

US-07-609-716-103
 ; Sequence 103, Application US/07609716
 ; Patent No. 5514581
 ; GENERAL INFORMATION:
 ; APPLICANT: Ferrari, Franco A.
 ; TITLE OF INVENTION: Functional Recombinantly Prepared
 ; TITLE OF INVENTION: Synthetic Protein Polymer
 ; NUMBER OF SEQUENCES: 118
 ; ADDRESS: Fiehr, Hobbach, Test, Albritton & Herbert
 ; STREET: Four Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 94111

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; SOFTWARE: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/07/609,716

FILING DATE: 06-NOV-1990

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Bertram I

REGISTRATION NUMBER: 20015

REFERENCE/DOCKET NUMBER: A-55186-3/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-07-609-716-103

Query Match 62.6%; Score 109; DB 1; Length 123;

Best Local Similarity 62.1%; Pred. No. 8.5e-07;
 Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLGTYGGLGGLGGLGGLGGLGGLGGLG 29
 |||||
 DB 35 GGVGVGGVGGVGGVGGVGGVGGVGGVGG 63

RESULT 3

US-08-475-411A-103

; Sequence 103, Application US/08475411A

; Patent No. 6140720X

; GENERAL INFORMATION:

; APPLICANT: Ferrari, Franco A.

; TITLE OF INVENTION: Functional Recombinantly Prepared

; TITLE OF INVENTION: Synthetic Protein Polymer

; NUMBER OF SEQUENCES: 119

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fiehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/475,411A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/609,716

; FILING DATE: 06-NOV-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/265,429

; FILING DATE: 07-OCT-1988

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/114,618

; FILING DATE: 29-OCT-1987

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 06/927,258

; FILING DATE: 04-NOV-1986

; ATTORNEY/AGENT INFORMATION:

; NAME: Frecartin, Richard F.

; REGISTRATION NUMBER: 31,801

; REFERENCE/DOCKET NUMBER: A-55186-9/REP/MTK

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-475-411A-103

COUNTRY: US
 ZIP: 94111
 COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; SOFTWARE: PC-DOS/MS-DOS
 ; CURRENT APPLICATION DATA: Patent Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/475,411A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/609,716

FILING DATE: 06-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/265,429

FILING DATE: 07-OCT-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/114,618

FILING DATE: 29-OCT-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/927,258

FILING DATE: 04-NOV-1986

ATTORNEY/AGENT INFORMATION:

NAME: Frecartin, Richard F.

REGISTRATION NUMBER: 31,801

REFERENCE/DOCKET NUMBER: A-55186-9/REP/MTK

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 103:

SEQUENCE CHARACTERISTICS:

LENGTH: 123 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-475-411A-103

Query Match 62.6%; Score 109; DB 4; Length 123;

Best Local Similarity 62.1%; Pred. No. 8.5e-07;
 Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLGTYGGLGGLGGLGGLGGLGGLGGLG 29
 |||||
 DB 35 GGVGVGGVGGVGGVGGVGGVGGVGGVGG 63

RESULT 4

US-08-478-029A-103

; Sequence 103, Application US/08478029A

; Patent No. 6184348

; GENERAL INFORMATION:

; APPLICANT: Ferrari, Franco A.

; TITLE OF INVENTION: Functional Recombinantly Prepared

; TITLE OF INVENTION: Synthetic Protein Polymer

; NUMBER OF SEQUENCES: 119

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fiehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/478,029A


```

: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/609,716
: FILING DATE: 06-NOV-1990
: PRIOR APPLICATION DATA: US 07/269,429
: FILING DATE: 09-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/114,618
: FILING DATE: 29-OCT-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 06/927,258
: FILING DATE: 04-NOV-1986
: ATTORNEY/AGENT INFORMATION:
: NAME: Treccartin, Richard F.
: REGISTRATION NUMBER: 31,801
: REFERENCE/DOCKET NUMBER: A-55186-8/RET/NTK
: TELEPHONE: 415-781-1989
: TELEFAX: 415-781-3249
: INFORMATION FOR SEQ ID NO: 103:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 123 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-478-029A-103

```

```

Query Match 62.6%; Score 109; DB 4; Length 123;
Best Local Similarity 62.1%; Pred. No. 8.5e-07;
Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 GGLGVLGGLGGLGGLGGLGGLG 29
DB 35 GGVGVGVGVGVGVGVGVGVGVGV 63

```

```

RESULT 5
US-07-609-716-104
: SEQUENCE 104, Application US/07609716
: Patent No. 5514581
: GENERAL INFORMATION:
: APPLICANT: Ferrari, Franco A.
: APPLICANT: Cappello, Joseph
: TITLE OF INVENTION: Functional Recombinantly Prepared
: TITLE OF INVENTION: Synthetic Protein Polymer
: NUMBER OF SEQUENCES: 118
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
: CITY: Fair Embarcadero Center, Suite 3400
: STATE: San Francisco
: COUNTRY: US
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 07/609,716
: FILING DATE: 06-NOV-1990
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Rowland, Bertam I
: REGISTRATION NUMBER: 20015
: REFERENCE/DOCKET NUMBER: A-55186-3/BIR
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-781-1989
: TELEFAX: 415-398-3245

```

```

: INFORMATION FOR SEQ ID NO: 104:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 159 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-07-609-716-104
Query Match 62.6%; Score 109; DB 1; Length 159;
Best Local Similarity 62.1%; Pred. No. 1.1e-06;
Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
QY 1 GGLGVLGGLGGLGGLGGLGGLG 29
DB 53 GGVGVGVGVGVGVGVGVGVGVGV 81

```

```

RESULT 6
US-08-475-411A-104
: SEQUENCE 104, Application US/08475411A
: Patent No. 6140872
: GENERAL INFORMATION:
: APPLICANT: Ferrari, Franco A.
: APPLICANT: Cappello, Joseph
: TITLE OF INVENTION: Functional Recombinantly Prepared
: TITLE OF INVENTION: Synthetic Protein Polymer
: NUMBER OF SEQUENCES: 119
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
: CITY: Fair Embarcadero Center, Suite 3400
: STATE: San Francisco
: COUNTRY: US
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 08/475,411A
: FILING DATE: 08-NOV-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/609,716
: FILING DATE: 06-NOV-1990
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/269,429
: FILING DATE: 09-NOV-1988
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/114,618
: FILING DATE: 29-OCT-1987
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 06/927,258
: FILING DATE: 04-NOV-1986
: ATTORNEY/AGENT INFORMATION:
: NAME: Treccartin, Richard F.
: REGISTRATION NUMBER: 31,801
: REFERENCE/DOCKET NUMBER: A-55186-9/RET/NTK
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-781-1989
: TELEFAX: 415-398-3249
: INFORMATION FOR SEQ ID NO: 104:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 159 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-08-475-411A-104

```

Query Match 62.6%; Score 109; DB 4; Length 159;

Best Local Similarity 62.1%; Pred. No. 1.1e-06;

Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLGVLGGVGGVGGVGGVGGVGGVGG 29

DB 53 GGVGGVGGVGGVGGVGGVGGVGGVGG 81

RESULT 7

US-08-478-029A-104

Sequence 104, Application US/08478029A

Patent No. 6184343

GENERAL INFORMATION:

APPLICANT: Ferrari, Franco A.

APPLICANT: Cappello, Joseph

TITLE OF INVENTION: Functional Recombinantly Prepared

TITLE OF INVENTION: Synthetic Protein Polymer

NUMBER OF SEQUENCES: 119

CORRESPONDENCE ADDRESS:

ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

STREET: Four Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: US

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,029A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/609,716

FILING DATE: 06-NOV-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/269,429

FILING DATE: 09-NOV-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/114,618

FILING DATE: 29-OCT-1987

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 06/927,258

FILING DATE: 04-NOV-1986

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: A-53186-8/PFT/WTX

REFERENCE/DOCKET NUMBER: A-53186-8/PFT/WTX

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 104:

SEQUENCE CHARACTERISTICS:

LENGTH: 159 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-478-029A-104

Query Match 62.6%; Score 109; DB 4; Length 159;

Best Local Similarity 62.1%; Pred. No. 1.1e-06;

Matches 18; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 GGLGVLGGVGGVGGVGGVGGVGGVGG 29

DB 53 GGVGGVGGVGGVGGVGGVGGVGGVGG 81

RESULT 8

US-08-963-168C-15

Sequence 15, Application US/08963168C

Patent No. 6127166

GENERAL INFORMATION:

APPLICANT: Bayley, Hagan

APPLICANT: Cao, Quiping

APPLICANT: Wang, Yunjaun

TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES

TITLE OF INVENTION: AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FAST50 for Windows Version 2.0

CURRENT APPLICATION DATA: US/08/963,168C

FILING DATE: 03-NOV-1997

ATTORNEY/AGENT INFORMATION:

NAME: Passe, Peter J.

REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 07917/059001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 111 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-963-168C-15

Query Match 55.2%; Score 96; DB 4; Length 111;

Best Local Similarity 50.0%; Pred. No. 2.6e-05;

Matches 17; Conservative 6; Mismatches 7; Indels 4; Gaps 1;

QY 1 GGLGVLGGVGGVGGVGGVGGVGGVGG 30

DB 78 GGGGAGGAGGAGGAGGAGGAGGAGGAGG 111

RESULT 9

US-08-963-168C-13

Sequence 13, Application US/08963168C

Patent No. 6127166

GENERAL INFORMATION:

APPLICANT: Bayley, Hagan

APPLICANT: Cao, Quiping

APPLICANT: Wang, Yunjaun

TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES

TITLE OF INVENTION: AND GENES ENCODING THEM

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

APPLICANT: Hayashi, Cheryl Y
TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
TITLE OF INVENTION: CODING THEREFOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 GATEHOUSE RD. SUITE 5002
CITY: FALLS CHURCH
STATE: VIRGINIA
ZIP: 22042
COUNTRY: UNITED STATES OF AMERICA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,928B
FILING DATE: 22-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28977
REFERENCE/DOCKET NUMBER: 1447-109P
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 745 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-010-928B-28

Query Match 53.4%; Score 93; DB 2; Length 745;
Best Local Similarity 62.1%; Pred. No. 0.00035;
Matches 18; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 GGLYGGGLGGLGGLGGLGGLGGLG 29
DB 702 GGAGFGAGFGGGLGAGRGAGRGAG 730

RESULT 13
US-09-010-928B-2
Sequence 2, Application US/09010928B
Patent No. 5994099
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph Y
APPLICANT: Hayashi, Cheryl Y
TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
TITLE OF INVENTION: CODING THEREFOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 GATEHOUSE RD. SUITE 5002
CITY: FALLS CHURCH
STATE: VIRGINIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,928B
FILING DATE: 22-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28977
REFERENCE/DOCKET NUMBER: 1447-109P
INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
LENGTH: 870 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-010-928B-2
Query Match 53.4%; Score 93; DB 2; Length 870;
Best Local Similarity 62.1%; Pred. No. 0.00041;
Matches 18; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
QY 1 GGLYGGGLGGLGGLGGLGGLGGLG 29
DB 828 GGAGFGAGFGGGLGAGRGAGRGAG 856
RESULT 14
US-09-010-928B-4
Sequence 4, Application US/09010928B
Patent No. 5994099
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph Y
APPLICANT: Hayashi, Cheryl Y
TITLE OF INVENTION: EXTREMELY ELASTIC SPIDER SILK PROTEIN AND DNA
TITLE OF INVENTION: CODING THEREFOR
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
STREET: 8110 GATEHOUSE RD. SUITE 5002
CITY: FALLS CHURCH
STATE: VIRGINIA
COUNTRY: UNITED STATES OF AMERICA
ZIP: 22042
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/010,928B
FILING DATE: 22-JAN-1998
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28977
REFERENCE/DOCKET NUMBER: 1447-109P
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 907 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-010-928B-4

Query Match 53.4%; Score 93; DB 2; Length 907;
Best Local Similarity 62.1%; Pred. No. 0.00042;
Matches 18; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

QY 1 GGLYGGGLGGLGGLGGLGGLGGLG 29
DB 728 GGAGFGAGFGGGLGAGRGAGRGAG 756

RESULT 15
US-08-963-169C-14
Sequence 14, Application US/08963168C
Patent No. 6127166
GENERAL INFORMATION:
APPLICANT: Bayley, Ragan
APPLICANT: Cao, Quiding
APPLICANT: Wang, Yurjaun

? TITLE OF INVENTION: MOLLUSCAN LIGAMENT POLYPEPTIDES
 ? TITLE OF INVENTION: AND GENS ENCODING THEM
 ? NUMBER OF SEQUENCES: 43
 ? CORRESPONDENCE ADDRESS:
 ? ADDRESSEE: Fish & Richardson P.C.
 ? STREET: 225 Franklin Street
 ? CITY: Boston
 ? STATE: MA
 ? COUNTRY: US
 ? ZIP: 02110-2804
 ? COMPUTER READABLE FORM:
 ? MEDIUM TYPE: Diskette
 ? COMPUTER: IBM Compatible
 ? OPERATING SYSTEM: Windows95
 ? SOFTWARE: FastSeq for Windows Version 2.0
 ? CURRENT APPLICATION DATA:
 ? APPLICATION NUMBER: US/08/963,168C
 ? FILING DATE: 03-NOV-1997
 ? ATTORNEY/AGENT INFORMATION:
 ? NAME: Fesse, Peter J.
 ? REGISTRATION NUMBER: 32,983
 ? REFERENCE/COCKET NUMBER: 07917/059001
 ? TELECOMMUNICATION INFORMATION:
 ? TELEPHONE: 617/542-5070
 ? TELEFAX: 617/542-8906
 ? TELEX: 200154
 ? INFORMATION FOR SEQ ID NO: 14:
 ? SEQUENCE CHARACTERISTICS:
 ? LENGTH: 106 amino acids
 ? TYPE: amino acid
 ? TOPOLOGY: linear
 ? WORDS: 1
 ? WORDS TYPE: Protein
 ? US-08-963-168C-14

Query Match 52.3%; Score 91; DB 4; Length 106;
 Best Local Similarity 50.0%; Pred. No. 9.6e-05;
 Matches 17; Conservative 6; Mismatches 7; Indels 4; Gaps 1;
 QY 1 GELAYGGLGGLTGGGLAYGRL---GIGGLAY 30
 DB 73 GMSMSMGFGMGGMGNGMGCGGCGGFGKGY 106

Search completed: April 24, 2001, 16:35:26
 time: 305 sec

A:Cross-references: EMBL:X15586; NID:g5771; PID:g5774
A>Note: This sequence was submitted to the EMBL Data Library, Jun-1999

A:Map position: 2
C:Superfamily: chorion class A protein pc292
F.I-21/Domain: signal sequence \$status predicted <SIG>
F.22-168/Product: chorion class B protein L11 \$status predicted <MAT>

Query Match 78.3%; Score 136; DB 2; Length 168;
Best Local Similarity 85.7%; Pred. No. 1e-08; Mismatches 0; Gaps 0;
Matches 24; Conservative 0; Indels 4;

QY 2 GLGVLGYGLGYSGLGYSGLGYSGLG 29
DB 25 GCGGGRGSGYGLGYSGLGYSGLGYSGLG 52

RESULT 5
S01420
Chorion class B protein (clone X2410) - silkworm (fragment)
Accession: S01420
Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Aug-1999
C:Accession: S01420
R:Tsitlikou, S.G.; Rodakis, G.C.; Alexopoulos, M.; Kafatos, P.C.; Ito, K.; Iatrou, K.
EMBO J. 2, 1845-1852, 1983
A:Title: Structural features of B family chorion sequences in the silkmoth Bombyx mori
A:Reference number: S01420; MUTD:84057707
A:Molecule type: mRNA
A:Accession: S01420
A:Cross-references: EMBL:X15586; NID:g5811; PID:CBA31322.1; PID:g5812
C:Superfamily: chorion class A protein pc292

Query Match 70.4%; Score 122.5; DB 2; Length 110;
Best Local Similarity 92.04%; Pred. No. 2e-07; Mismatches 0; Gaps 1;
Matches 23; Conservative 0; Indels 1;

QY 5 YGSLGYGLGYSGLGYSGLGYSGLG 29
DB 1 YGSLGYGLGYSGLGYSGLGYSGLG 24

RESULT 6
EXMS
elastin precursor - mouse
Alternate names: tropoelastin
N:Alternate names: tropoelastin
C:Date: 18-Aug-1995 #sequence_revision 16-Aug-1996 #text_change 22-Jun-1999
C:Accession: A55721
R:Widmer, K.S.; Sechler, J.L.; Boyd, C.D.; Passmore, H.C.
Genomics 23, 125-131, 1994
A:Title: Use of an intron length polymorphism to localize the tropoelastin gene to chromosome 23
A:Reference number: A55721; MUTD:95130669
A:Accession: A55721
A:Molecule type: mRNA
A:Residues: 1-860 <GB>
A:Cross-references: EMBL:U08210; NID:g473273; PID:AAA80155.1; PID:g473274
C:Accession: U08210
A:Map position: 5
C:Superfamily: elastin
C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxylysine
F.I-27/Domain: signal sequence \$status predicted <SIG>
F.26-860/Product: elastin \$status predicted <MAT>
F.850-855/Disulfide bonds: \$status predicted

Query Match 70.1%; Score 122; DB 1; Length 860;
Best Local Similarity 82.8%; Pred. No. 1e-06; Mismatches 5; Indels 0;
Matches 24; Conservative 0; Gaps 0;

OY 1 GLGVLGYGLGYSGLGYSGLGYSGLG 29

db 738 GGLGAGGLGAGGLGAGGLGAGGLGAGGLG 766

RESULTS

GenBank accession number: AF090872
 Title: Evolution of two major chorion multigene families as inferred from cloned cDNA
 Accession: AF090872
 Date: 28-Feb-1999
 Revision: 28-Feb-1999
 Author: R. J. Jones, C. W. Rosenthal, N. J. Rodakis, G. C. Kafatos, F. C. Kafatos
 Journal: Development
 Volume: 127
 Issue: 1
 Pages: 1317-1332
 Year: 1999

A: Molecule type: mRNA
A: Residues: 1-171 <ON>
A: Cross-references: GB:J01162; ND:J5630; PIDN:CAA23419.1; PID:J5631
C: Comment: This protein is one of many from the eggshell of the polyripemus moth.
C: Superfamily: chorion class A protein pc292
C: Keywords: egg shell; structural protein
F: 1-12/Domain; signal sequence #status predicted <SIG>
F: 23-171/Product; chorion class B protein pc401 #status predicted <MAT>

```

Query Match      68.4%; Score 119; DB 3; Length 171;
Best local Similarity 75.9%; Pred. No. 7.5e-07;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;
QY      1 GGLATGAGTGGTGGTGGTGGTGGTGGTGG 29
DB      136 GGLATGAGTGGTGGTGGTGGTGGTGGTGG 164

```

RESULT 9
JMA092 class A protein pC292 precursor - polyphemus moth (fragment)
C.Species: Anthraea polyphemus (polyphemus moth)
C.Date: 31-Aug-1980 #sequence_revision 31-Aug-1980 #text_change 15-Oct-1999
C.Accession: A03336
R.Titelius, S.G.; Regier, J.C.; Kafatos, F.C.
Nucleic Acids Res. 8, 1987-1997
A.Reference number: A03336
A.Reference number: A03336; MIM:81033859

A:Residues: 1-119 <N>
A:Cross-references: GB:P00078; NID:P5632; PIDD:CAA3420.1; PID:p5633
A:Note: partial direct sequencing of the protein revealed one difference from the above sequence
C:Superfamily: chorion class A protein p292
C:Keywords: egg shell; structural protein
F:1-6/Domains: signal sequence (fragment) #status predicted <S>
F:7-119/Product: chorion class A protein p292 #status predicted <A>

A1Note: partial direct sequencing of the protein revealed one difference from the
 C3proteinfamily: chorion class A protein pc232
 C3keywords: egg shell; structural protein
 F7.7-119/Domain: signal sequences (fragment) #status predicted <SIG>
 F7.7-119/Product: chorion class A protein pc232 #status predicted <MAT>
 Query Match 66.4% Score 115.5 Da 1 Length 119:
 Best Local Similarity 68.6% Pval 1.3e-06

keratin, scale - chicken
N:Alternate names: beta keratin
C:Species: Gallus gallus (chicken)
CDate: 17-Mar-1987 #sequence revision 17-Mar-1987 #text_change 22-Jun-1999
CAccession: A02831; S60136; I50168
R:Gregg, K.; Milton, S.D.; Rogers, G.E.
EMBO J. 3, 175-178, 1984

EMEO J. 3, 175-178, 1984

Query Match 59.5%; Score 103.5; DB 2; Length 89;
Best Local Similarity 47.1%; Pred. No. 2e-05;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 24, 2001, 16:42:52 ; Search time 44.68 seconds
(without alignments)
22.898 Million cell updates/sec

Title: US-09-340-736-8

Sequence: 1 GGLGVLGGLGGLGGLGGLGGLGGLY 30

Scoring table: BLCSUM62

Gapop 10.0 , Gapext 0.5

Archived: 93435 seqs, 34253486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 04

Listing first 45 summaries

Database: SwissProt_39.*

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	160	92.0	119	1	LAMP_PETMA
2	146	83.9	139	1	LAMP_PETMA
3	138	78.2	168	1	CHB1_BOMMO
4	136	78.2	168	1	CHB1_BOMMO
5	122.5	70.4	110	1	CHB6_BOMMO
6	122	70.1	860	1	ELIS_MOUSE
7	122	70.1	864	1	ELIS_RAT
8	119	68.4	171	1	CHB4_ANTPO
9	118.5	68.1	161	1	CHB2_BOMMO
10	115.5	66.4	119	1	CHAL_ANTPO
11	109.5	62.9	154	1	KHSC_CHICK
12	104.5	60.1	91	1	CHB3_BOMMO
13	104.5	60.1	91	1	CHB3_BOMMO
14	104.5	59.9	643	1	KHSC_HUMAN
15	102	58.6	152	1	CUG4_LOOMI
16	101.5	58.3	512	1	K2C3_XENLA
17	100	57.5	79	1	KHRA_RABIT
18	98	56.3	1380	1	DX93_MOUSE
19	97	55.7	168	1	GRP2_SORBI
20	96.5	55.5	126	1	CHB7_BOMMO
21	95.5	54.9	750	1	ELIS_CHICK
22	94.5	54.3	145	1	CUG5_LOOMI
23	94	54.0	127	1	KHCL_CHICK
24	93	53.4	129	1	CHB5_MOUSE
25	92	52.9	195	1	KHRA_XENLA
26	92	52.9	373	1	KHRA_XENLA
27	92	52.6	385	1	KHRA_XENLA
28	91.5	52.3	521	1	OCLN_MOUSE
29	91	52.3	627	1	K2C1_MOUSE
30	91	52.0	157	1	CUG3_LOOMI
31	90.5	51.7	169	1	GR10_BRANA
32	90	51.7	208	1	AC20_TENMO
33	90	51.7	208	1	AC20_TENMO

ALIGNMENTS

RESULT 1	LAMP_PETMA	STANDARD	PRN: 119 AA.
ID	LAMP_PETMA	AC	P33577:
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DT	01-FEB-1994 (Rel. 28, Last annotation update)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DE	LAMPIN 1.8-10 PRECURSOR (CARLAGE MATRIX PROTEIN).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (Sea lamprey).		
CC	Petromyzon marinus (

Query Match 92.0% Score 160; DB 1; Length 119;
Best Local Similarity 93.3%; Fred. No. 4.6e-10;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

[illegible]

```

Query Match      68.4%; Score 119; DB 1; Length 171;
Best Local Similarity 75.9%; Pred. 6.2e-06;
Matches 22; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY      1  GGLAYGGLGGLGGLGGLGGLGGLG 29
          ||||| ||| | | | | | | |
DB      136  GGLAYGGLGGLGGLGGLGGLGGLG 164

```

```

RESULT 9
CHB2.BOMO STANDARD: PRT: 161 AA.
ID CHB2.BOMO
AC P08828;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DE CHORION CLASS B PROTEIN L12 PRECURSOR.
OS Bombyx mori. (Silk moth).
OC Arthropoda; Insecta;
OC Bombyx mori; Metazoa; Arthropoda; Insecta;
OC Bombycidae; Bombyx;
OC NCBI_TaxID=7091;
RN RE
RE SEQUENCE FROM N.A.
MC MEDLINE=703;
MC MEDLINE=7060979; PubMed=3023635;
MC Sporel N.T., Nguyen H.T., Kafatos F.C.;
MC "Gene regulation and evolution in the chorion locus of Bombyx mori."
MC "Structural and developmental characterization of four eggshell genes
MC and their flanking DNA regions."
MC PLoS Biol. 1(1):E169; 2003.
CC -!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
CC -!- SILK MOTH.
CC -!- SIMILARITY: MEMBER OF THE BETA-BRANCH OF CHORION PROTEIN TO WHICH
CC BELONG CLASSES B, CB AND HCB.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X15557; CAA33564.1;
CC PIR: D24255;
CC InterPro: IPR002635;
CC Pfam: PF01723; Chorion; Repeat; Multigene family; Signal.
CC Eggshell; Chorion; Repeat; Multigene family; Signal.
CC CHAIN 22 161 CHORION CLASS B PROTEIN L12.
CC DOMAIN 22 52 LEFT ARM.
CC DOMAIN 53 121 CENTRAL DOMAIN.
CC DOMAIN 122 161 RIGHT ARM (GLY-RICH TANDEM REPEATS).
CC DOMAIN 30 44 3 X 5 AA TANDEM REPEATS OF G-Y-G-L.
CC SEQUENCE 161 AA: 15123 MW; 782FDCA9FC07FA0B CRC64;
CC
CC Query Match 68.1%; Score 118.5; DB 1; Length 161;
CC Best Local Similarity 82.1%; Pred. No. 6; e=0.0;
CC Matches 23; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
CC
CC QY 2 GLGYGLGYGLGYGLGYGLGYGLGYGLGY 29
CC |||||
CC DB 23 GCGGCGGGLGYGLGYGLGYGLGYGLGY 49
CC
CC RESULT 10
CHALANPO STANDARD: PRT: 119 AA.
ID CHALANPO
AC P08828;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE CHORION CLASS A PROTEIN PC292 PRECURSOR (FRAGMENT).
OS Antherea polyphemus (Polyphemus moth).
OC Arthropoda; Insecta;
OC Antherea polyphemus; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Bombycoidea; Saturniidae; Saturniinae; Antherea.
CC NCBI_TaxID=7120;
CC

```

```

SEQUENCE FROM N.A.
MEDLINE=81053859; PubMed=7431333;
Tsitolou S.G., Reider J.C., Kafatos F.C.;
"Chorion protein of the A family."
Nucleic Acids Res. 8:1987-1997(1980).
-!- FUNCTION: THIS PROTEIN IS ONE OF MANY FROM THE EGGSHELL OF THE
-!- SILK MOTH.
-!- SIMILARITY: MEMBER OF THE ALPHA-BRANCH OF CHORION PROTEIN TO WHICH
-!- BELONG CLASSES A, CA AND HCA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: J01159; AAA27780.1;
CC EMBL: V00078; CAA23420.1;
CC PIR: A03367; JAA092;
CC Eggshell; Chorion; Repeat; Multigene family; Signal.
CC NCBI_TaxID=7091;
CC CHAIN 7 119 CHORION CLASS A PROTEIN PC292.
CC DOMAIN 7 53 LEFT ARM.
CC DOMAIN 54 102 CENTRAL DOMAIN.
CC DOMAIN 103 119 RIGHT ARM.
CC REPEAT 28 32
CC REPEAT 33 37
CC REPEAT 38 42
CC REPEAT 113 116
CC SEQUENCE 119 AA: 11091 MW; 65E01B7C0F2A2AA CRC64;
CC
CC Query Match 66.4%; Score 115.5; DB 1; Length 119;
CC Best Local Similarity 68.6%; Pred. No. 1e-05;
CC Matches 24; Conservative 0; Mismatches 6; Indels 5; Gaps 1;
CC
CC QY 1 GLGYGLGYGLGYGLGYGLGYGLGYGLGY 30
CC |||||
CC DB 10 GGLGLGLAPACGCGGLGYGLGYGLGYGLGY 44
CC
CC RESULT 11
KRSC.CHICK STANDARD: PRT: 154 AA.
ID KRSC.CHICK
AC P04459;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DE KERATIN, SCALP (S-KER)
DE KERATIN, SCALP (S-KER)
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CC NCBI_TaxID=9031;
CC
CC SEQUENCE FROM N.A.
MEDLINE=84158528; PubMed=5200321;
Rogers G.E.;
"Comparison of genomic coding sequences for feather and scale
keratins: structural and evolutionary implications."
EMBO J. 3:175-178(1984).
-!- MISCELLANEOUS: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER)
-!- ARE A COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

Result	Query No.	Score	Match	Query Length	DB ID	Description
1	150	92.0	138	13	Q9VHE0	Q9VHE0 petromyzon
2	146	83.9	119	13	Q9VGD0	Q9VGD0 petromyzon
3	146	83.9	138	13	Q9VSE4	Q9VSE4 petromyzon
4	146	83.9	142	13	Q9VSE2	Q9VSE2 petromyzon
5	120	89.0	142	5	Q9U517	Q9U517 manduca sex
6	116	86.7	33	3	Q9Z711	Q9Z711 antleraeae p
7	116	86.7	33	3	Q9Z711	Q9Z711 antleraeae p
8	109.5	82.9	56	13	Q87483	Q87483 gallus gall
9	109.5	82.9	99.4	5	Q9V671	Q9V671 mytilus edu
10	108	82.1	181	5	Q9VWK5	Q9VWK5 dirosophila
11	106.5	61.2	259	5	O02049	O02049 caenorhabdi
12	105.5	60.6	77	11	Q08621	Q08621 mus musculu
13	105.5	60.6	78	11	Q08048	Q08048 mus musculu
14	104	59.8	85	11	Q08636	Q08636 mus musculu
15	103.5	59.5	287	5	Q17201	Q17201 bombyx mori
16	103.5	59.5	287	5	Q17201	Q17201 bombyx mori
17	103.5	59.5	303	5	Q17201	Q17201 dirosophila
18	102	58.6	945	5	Q9W4V1	Q9W4V1 dirosophila
19	102	58.6	1894	5	Q9NHW2	Q9NHW2 nephila nas

Query Match 92.0%; Score 160; DB 13; Length 138;
Best Local Similarity 93.3%; Pred. NO. 2.3e-11;
Matches 28; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Query Match 92.08; Score 150; DB 13; Length 138;

QY 1 GGLGYGGLGYGGLGYGGLGYGGLGY 30
 |||||
 DQ 41 GGLGYGGLGYGGLGYGGLGYGGLGY 70
 |||||

[illegible]

```

RESULT 2
ID Q9PSW3 PRELIMINARY: PRT: 119 AA.
AC Q9PSW3
DT 01-MAY-2000 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE LAMPIN L-0.9-10.
GN L-0.9.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Youson J.H., Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97499.1; JOINED.
RA EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 119 AA; 11145 MW; BE608EE263DF03ED CRC64;

Query Match
Best Local Similarity 83.9%; Score 146; DB 13; Length 119;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 3
ID Q9PSW3 PRELIMINARY: PRT: 138 AA.
AC Q9PSW3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE LAMPIN L-0.8-12.
GN L-0.8.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Youson J.H., Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97499.1; JOINED.
RA EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 138 AA; 13094 MW; AB2F35D0D25D7580 CRC64;

Query Match
Best Local Similarity 86.7%; Pred. No. 8.1e-10;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 4
ID Q9PSW2 PRELIMINARY: PRT: 138 AA.
AC Q9PSW2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE LAMPIN L-0.9-12.
GN L-0.9-12.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97501.1; JOINED.
RA EMBL: AF097758; AAC97501.1; JOINED.
DR EMBL: AF097758; AAC97501.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 138 AA; 13144 MW; A32F35D0C74924D0 CRC64;

Query Match
Best Local Similarity 83.9%; Score 146; DB 13; Length 138;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 5
ID Q9US17 PRELIMINARY: PRT: 142 AA.
AC Q9US17
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE PUTATIVE CUTICLE PROTEIN.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Spingidae; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Roberts H.M., Martos R., Sears C.R., Todres E.Z., Walden K.K.O.,
RA Nord J.B.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF117571; AAF16693.1; JOINED.
DR EMBL: AF117571; AAF16693.1; JOINED.
SQ SEQUENCE 142 AA; 14112 MW; 17230D5D5564DE80 CRC64;

```

```

RESULT 2
ID Q9PSW3 PRELIMINARY: PRT: 119 AA.
AC Q9PSW3
DT 01-MAY-2000 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE LAMPIN L-0.9-10.
GN L-0.9.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Youson J.H., Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97499.1; JOINED.
RA EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 119 AA; 11145 MW; BE608EE263DF03ED CRC64;

Query Match
Best Local Similarity 83.9%; Score 146; DB 13; Length 119;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 3
ID Q9PSW3 PRELIMINARY: PRT: 138 AA.
AC Q9PSW3
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE LAMPIN L-0.8-12.
GN L-0.8.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97499.1; JOINED.
RA EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097757; AAC97499.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 138 AA; 13094 MW; AB2F35D0D25D7580 CRC64;

Query Match
Best Local Similarity 86.7%; Pred. No. 8.1e-10;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 4
ID Q9PSW2 PRELIMINARY: PRT: 138 AA.
AC Q9PSW2
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE LAMPIN L-0.9-12.
GN L-0.9-12.
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hypercoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF097757; AAC97501.1; JOINED.
RA EMBL: AF097758; AAC97501.1; JOINED.
DR EMBL: AF097758; AAC97501.1; JOINED.
DR EMBL: AF097756; AAC97498.1; JOINED.
SQ SEQUENCE 138 AA; 13144 MW; A32F35D0C74924D0 CRC64;

Query Match
Best Local Similarity 83.9%; Score 146; DB 13; Length 138;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 GGLGSGGLGGLGGLGGLGGLGGLG 30
DB 41 GGLGSGGLGGLGGLGGLGGLGGLG 70

RESULT 5
ID Q9US17 PRELIMINARY: PRT: 142 AA.
AC Q9US17
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE PUTATIVE CUTICLE PROTEIN.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Spingidae; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP MEDLINE=9123269; PubMed=7678258;
RA Roberts H.M., Martos R., Sears C.R., Todres E.Z., Walden K.K.O.,
RA Nord J.B.;
RA Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF117571; AAF16693.1; JOINED.
DR EMBL: AF117571; AAF16693.1; JOINED.
SQ SEQUENCE 142 AA; 14112 MW; 17230D5D5564DE80 CRC64;

```

Query Match 69.0%; Score 120; DB 5; Length 142;
 Best Local Similarity 77.4%; Pred. No. 6.5e-07;
 Matches 24; Conservative 1; Mismatches 2; Indels 4; Gaps 2;

QY 2 GUGYGGGAYG--GUGYGGGAYGUGYGGY 30
 DB 29 GUGYGGGAYG--GUGYGGGAYGUGYGGY 57

RESULT 6
 ID Q17071 PRELIMINARY; PRT: 33 AA.
 AC Q17071.
 DT 01-NOV-1996 (TREMELrel. 01, Created)
 DT 01-NOV-1996 (TREMELrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE SILKMOH (A. POLYPHEMUS) CHORION PROTEIN (FRAGMENT).
 OS Silkworm, Lepidoptera; Tracheata; Hexapoda; Insecta;
 OC Bombycidae; Saturniidae; Saturniinae; Antheraea.
 CC Bombycidae; Saturniidae; Saturniinae; Antheraea.
 CB NCBI_TaxID=7120;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=80090072; PubMed=519771;
 RA Jones C.W., Rosenthal N., Rodakis G.C., Kafatos F.C.;
 RT "Evolution of two major chorion multigene families as inferred from
 RT cloned cDNA and protein sequences";
 RU Cell 18:1317-1332(1979).
 DU 18:1317-1332(1979).
 KW Chorionic, Multigene family.
 FT NON_TER 33 33
 SQ SEQUENCE 33 AA; 2570 MW; 60D3208D2D16288B CRC64;

Query Match 66.7%; Score 116; DB 5; Length 33;
 Best Local Similarity 78.6%; Pred. No. 3.7e-07;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 GUGYGGGAYG--GUGYGGGAYGUGYGGY 29
 DB 4 GUGYGGGAYG--GUGYGGGAYGUGYGGY 31

RESULT 7
 ID P82170 PRELIMINARY; PRT: 62 AA.
 AC P82170.
 DT 01-MAY-2000 (TREMELrel. 13, Created)
 DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMELrel. 13, Last annotation update)
 DE CUTICLE PROTEIN 5 (LWKP6)
 OS Locusta migratoria (Migratory locust).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Orthopteroidea; Orthoptera; Caelifera;
 OC Acridoidea; Acridoidea; Acrididae; Locusta.
 CB NCBI_TaxID=7004;
 RN [1]
 RP SEQUENCE.
 RX SUBMITTER (NOV-1999) TO THE SWISS-PROT DATA BANK.
 RA Submitter (NOV-1999) TO THE SWISS-PROT DATA BANK.
 CC -1- FUNCTION: COMPONENT OF THE CUTICLE OF MIGRATORY LOCUST WHICH
 CC CONTAINS MORE THAN 100 DIFFERENT STRUCTURAL PROTEINS.
 KW Structural protein; Cuticle.
 FT NON_TER 62 62
 SQ SEQUENCE 62 AA; 6438 MW; F3762E1E8F17D95B CRC64;

Query Match 66.7%; Score 116; DB 5; Length 62;
 Best Local Similarity 63.9%; Pred. No. 7.3e-07;
 Matches 23; Conservative 0; Mismatches 7; Indels 6; Gaps 1;

QY 1 GGLY-----GGLYGGGAYGUGYGGYGGY 30
 DB 21 GGLYGGGAYGUGYGGGAYGUGYGGYGGY 56

RESULT 8
 ID P87483 PRELIMINARY; PRT: 66 AA.
 AC P87483.
 DT 01-MAY-1997 (TREMELrel. 03, Created)
 DT 01-MAY-1997 (TREMELrel. 03, Last sequence update)
 DT 01-NOV-1998 (TREMELrel. 08, Last annotation update)
 DE KERATIN (FRAGMENT).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 CB NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91278000; PubMed=2570818;
 RA White S.D.
 RT Isolation and characterization of keratin mRNA from the scale
 RT epidermis of the embryonic chick.;
 RL Biochim. Biophys. Acta 824:201-208(1985).
 DR EMBL; M25642; AAA48931.1;
 KW Keratin.
 FT NON_TER 1 1
 SQ SEQUENCE 66 AA; 5813 MW; EBC3A87D99918F25 CRC64;

Query Match 52.9%; Score 109.5; DB 13; Length 66;
 Best Local Similarity 76.9%; Pred. No. 7.1e-06;
 Matches 30; Conservative 0; Mismatches 0; Indels 9; Gaps 6;

QY 1 GG--LGYGGL--GYGG--LGYGGL--GY 30
 DB 1 GSSSLGYGGLGYGGSSLYGGLGYGGSSLYGGLGY 39

RESULT 9
 ID Q76271 PRELIMINARY; PRT: 904 AA.
 AC Q76271.
 DT 01-NOV-1998 (TREMELrel. 08, Created)
 DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
 DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
 DE NONGRADIENT BYSSAL PRECURSOR.
 OS Mytilus edulis (Blue mussel).
 OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;
 OC Mytiloidea; Mytilidae; Mytilus.
 CB NCBI_TaxID=6550;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98393676; PubMed=9724735;
 RA Qin X.X., Waite J.H.;
 RT "A potential mediator of collagenous block copolymer gradients in
 RT mussel byssal threads";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:10517-10522(1998).
 DR EMBL; AF043944; AAC33847.1;
 DR INTERPRO: IPR000087;
 DR PFAM: PF01391; Collagen; 7
 SQ SEQUENCE 904 AA; 77683 MW; 5529135651AD4C40 CRC64;

Query Match 62.9%; Score 109.5; DB 5; Length 904;
 Best Local Similarity 82.8%; Pred. No. 7.1e-05;
 Matches 24; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 GGLYGGGAYGUGYGGGAYGUGYGGYGGY 29
 DB 140 GGLYGGGAYGUGYGGGAYGUGYGGYGGY 167

```

ID AC Q02049 PRELIMINARY; PROT; 259 AA.
OC Q02049
DI DT 01-JUL-1997 (T=EMBLrel_04, Created)
DI DT 01-JUL-1997 (T=EMBLrel_04, Last sequence update)
DI DT 01-OCT-2000 (T=EMBLrel_15, Last annotation update)
DE COXKID T20B6.
DE DT T20B6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodirinae; Caenorhabditis.
NCBI_TextID:6239;
RX (1)
RX (2)
RX (3)
RX (4)
RX (5)
RX (6)
RX (7)
RX (8)
RX (9)
RX (10)
RX (11)
RX (12)
RX (13)
RX (14)
RX (15)
RX (16)
RX (17)
RX (18)
RX (19)
RX (20)
RX (21)
RX (22)
RX (23)
RX (24)
RX (25)
RX (26)
RX (27)
RX (28)
RX (29)
RX (30)
RX (31)
RX (32)
RX (33)
RX (34)
RX (35)
RX (36)
RX (37)
RX (38)
RX (39)
RX (40)
RX (41)
RX (42)
RX (43)
RX (44)
RX (45)
RX (46)
RX (47)
RX (48)
RX (49)
RX (50)
RX (51)
RX (52)
RX (53)
RX (54)
RX (55)
RX (56)
RX (57)
RX (58)
RX (59)
RX (60)
RX (61)
RX (62)
RX (63)
RX (64)
RX (65)
RX (66)
RX (67)
RX (68)
RX (69)
RX (70)
RX (71)
RX (72)
RX (73)
RX (74)
RX (75)
RX (76)
RX (77)
RX (78)
RX (79)
RX (80)
RX (81)
RX (82)
RX (83)
RX (84)
RX (85)
RX (86)
RX (87)
RX (88)
RX (89)
RX (90)
RX (91)
RX (92)
RX (93)
RX (94)
RX (95)
RX (96)
RX (97)
RX (98)
RX (99)
RX (100)
RX (101)
RX (102)
RX (103)
RX (104)
RX (105)
RX (106)
RX (107)
RX (108)
RX (109)
RX (110)
RX (111)
RX (112)
RX (113)
RX (114)
RX (115)
RX (116)
RX (117)
RX (118)
RX (119)
RX (120)
RX (121)
RX (122)
RX (123)
RX (124)
RX (125)
RX (126)
RX (127)
RX (128)
RX (129)
RX (130)
RX (131)
RX (132)
RX (133)
RX (134)
RX (135)
RX (136)
RX (137)
RX (138)
RX (139)
RX (140)
RX (141)
RX (142)
RX (143)
RX (144)
RX (145)
RX (146)
RX (147)
RX (148)
RX (149)
RX (150)
RX (151)
RX (152)
RX (153)
RX (154)
RX (155)
RX (156)
RX (157)
RX (158)
RX (159)
RX (160)
RX (161)
RX (162)
RX (163)
RX (164)
RX (165)
RX (166)
RX (167)
RX (168)
RX (169)
RX (170)
RX (171)
RX (172)
RX (173)
RX (174)
RX (175)
RX (176)
RX (177)
RX (178)
RX (179)
RX (180)
RX (181)
RX (182)
RX (183)
RX (184)
RX (185)
RX (186)
RX (187)
RX (188)
RX (189)
RX (190)
RX (191)
RX (192)
RX (193)
RX (194)
RX (195)
RX (196)
RX (197)
RX (198)
RX (199)
RX (200)
RX (201)
RX (202)
RX (203)
RX (204)
RX (205)
RX (206)
RX (207)
RX (208)
RX (209)
RX (210)
RX (211)
RX (212)
RX (213)
RX (214)
RX (215)
RX (216)
RX (217)
RX (218)
RX (219)
RX (220)
RX (221)
RX (222)
RX (223)
RX (224)
RX (225)
RX (226)
RX (227)
RX (228)
RX (229)
RX (230)
RX (231)
RX (232)
RX (233)
RX (234)
RX (235)
RX (236)
RX (237)
RX (238)
RX (239)
RX (240)
RX (241)
RX (242)
RX (243)
RX (244)
RX (245)
RX (246)
RX (247)
RX (248)
RX (249)
RX (250)
RX (251)
RX (252)
RX (253)
RX (254)
RX (255)
RX (256)
RX (257)
RX (258)
RX (259)
RX (260)
RX (261)
RX (262)
RX (263)
RX (264)
RX (265)
RX (266)
RX (267)
RX (268)
RX (269)
RX (270)
RX (271)
RX (272)
RX (273)
RX (274)
RX (275)
RX (276)
RX (277)
RX (278)
RX (279)
RX (280)
RX (281)
RX (282)
RX (283)
RX (284)
RX (285)
RX (286)
RX (287)
RX (288)
RX (289)
RX (290)
RX (291)
RX (292)
RX (293)
RX (294)
RX (295)
RX (296)
RX (297)
RX (298)
RX (299)
RX (300)
RX (301)
RX (302)
RX (303)
RX (304)
RX (305)
RX (306)
RX (307)
RX (308)
RX (309)
RX (310)
RX (311)
RX (312)
RX (313)
RX (314)
RX (315)
RX (316)
RX (317)
RX (318)
RX (319)
RX (320)
RX (321)
RX (322)
RX (323)
RX (324)
RX (325)
RX (326)
RX (327)
RX (328)
RX (329)
RX (330)
RX (331)
RX (332)
RX (333)
RX (334)
RX (335)
RX (336)
RX (337)
RX (338)
RX (339)
RX (340)
RX (341)
RX (342)
RX (343)
RX (344)
RX (345)
RX (346)
RX (347)
RX (348)
RX (349)
RX (350)
RX (351)
RX (352)
RX (353)
RX (354)
RX (355)
RX (356)
RX (357)
RX (358)
RX (359)
RX (360)
RX (361)
RX (362)
RX (363)
RX (364)
RX (365)
RX (366)
RX (367)
RX (368)
RX (369)
RX (370)
RX (371)
RX (372)
RX (373)
RX (374)
RX (375)
RX (376)
RX (377)
RX (378)
RX (379)
RX (380)
RX (381)
RX (382)
RX (383)
RX (384)
RX (385)
RX (386)
RX (387)
RX (388)
RX (389)
RX (390)
RX (391)
RX (392)
RX (393)
RX (394)
RX (395)
RX (396)
RX (397)
RX (398)
RX (399)
RX (400)
RX (401)
RX (402)
RX (403)
RX (404)
RX (405)
RX (406)
RX (407)
RX (408)
RX (409)
RX (410)
RX (411)
RX (412)
RX (413)
RX (414)
RX (415)
RX (416)
RX (417)
RX (418)
RX (419)
RX (420)
RX (421)
RX (422)
RX (423)
RX (424)
RX (425)
RX (426)
RX (427)
RX (428)
RX (429)
RX (430)
RX (431)
RX (432)
RX (433)
RX (434)
RX (435)
RX (436)
RX (437)
RX (438)
RX (439)
RX (440)
RX (441)
RX (442)
RX (443)
RX (444)
RX (445)
RX (446)
RX (447)
RX (448)
RX (449)
RX (450)
RX (451)
RX (452)
RX (453)
RX (454)
RX (455)
RX (456)
RX (457)
RX (458)
RX (459)
RX (460)
RX (461)
RX (462)
RX (463)
RX (464)
RX (465)
RX (466)
RX (467)
RX (468)
RX (469)
RX (470)
RX (471)
RX (472)
RX (473)
RX (474)
RX (475)
RX (476)
RX (477)
RX (478)
RX (479)
RX (480)
RX (481)
RX (482)
RX (483)
RX (484)
RX (485)
RX (486)
RX (487)
RX (488)
RX (489)
RX (490)
RX (491)
RX (492)
RX (493)
RX (494)
RX (495)
RX (496)
RX (497)
RX (498)
RX (499)
RX (500)
RX (501)
RX (502)
RX (503)
RX (504)
RX (505)
RX (506)
RX (507)
RX (508)
RX (509)
RX (510)
RX (511)
RX (512)
RX (513)
RX (514)
RX (515)
RX (516)
RX (517)
RX (518)
RX (519)
RX (520)
RX (521)
RX (522)
RX (523)
RX (524)
RX (525)
RX (526)
RX (527)
RX (528)
RX (529)
RX (530)
RX (531)
RX (532)
RX (533)
RX (534)
RX (535)
RX (536)
RX (537)
RX (538)
RX (539)
RX (540)
RX (541)
RX (542)
RX (543)
RX (544)
RX (545)
RX (546)
RX (547)
RX (548)
RX (549)
RX (550)
RX (551)
RX (552)
RX (553)
RX (554)
RX (555)
RX (556)
RX (557)
RX (558)
RX (559)
RX (560)
RX (561)
RX (562)
RX (563)
RX (564)
RX (565)
RX (566)
RX (567)
RX (568)
RX (569)
RX (570)
RX (571)
RX (5
```

QY 1 GGLGYGGLGYGGLGYGGLGYGGLGY 30

db 72 GGPFGGPGFGGPGFGGPGFGGPGFG 101

RESULT 11

640200

SEQUENCE	54 AA;	5729 MW;	2E7880E8822C1787 CRC64;
SQ			

SEQ
 SEQUENCE 89 AA; 9391 MW; A40C35A084858494 CRC64;
 EMBL; U04837; AAB04833.1;

```

search completed:
Job time: 431 sec

```
